

AD-A171 575

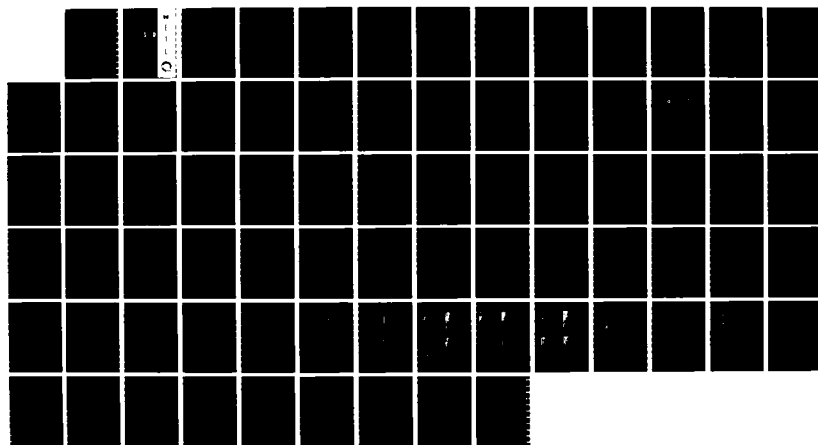
DEFENSE MAPPING AGENCY ADVANCED RASTER-TO-VECTOR  
BENCHMARK TESTING(U) BATTELLE COLUMBUS DIV WASHINGTON  
DC D D SELDEN JUN 86 ETL-0420 DAAG29-81-D-0100

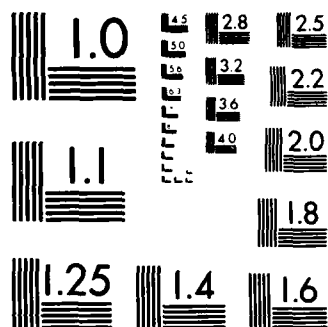
1/1

UNCLASSIFIED

F/G 8/2

NL





/

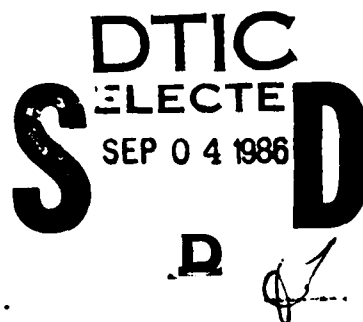
ETL - 0420

AD-A171 575

②

Defense Mapping Agency  
advanced raster - to - vector  
benchmark testing

David D. Selden



BATTELLE  
Columbus Laboratories  
Washington, D.C., Operations  
2030 M Street, NW  
Washington, D.C. 20036

June 1986

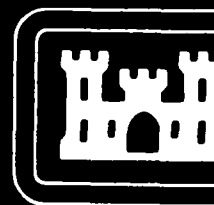
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

DTIC FILE COPY

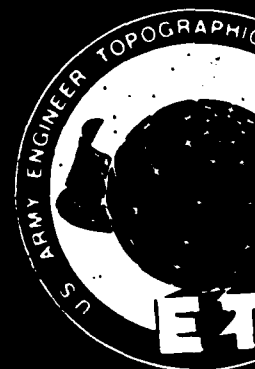
Prepared for  
U.S. ARMY CORPS OF ENGINEERS  
ENGINEER TOPOGRAPHIC LABORATORIES  
FORT BELVOIR, VIRGINIA 22060 - 5546

Monitored by  
DEFENSE MAPPING AGENCY  
HYDROGRAPHIC/TOPOGRAPHIC CENTER  
WASHINGTON, D.C. 20315 - 0030

86 9 3 113



E  
T  
L



Destroy this report when no longer needed.  
Do not return it to the originator.

---

The findings in this report are not to be construed as an official  
Department of the Army position unless so designated by other  
authorized documents.

---

The citation in this report of trade names of commercially available  
products does not constitute official endorsement or approval of the  
use of such products.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ETL-0420	2. GOVT ACCESSION NO. ADA171575	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DEFENSE MAPPING AGENCY ADVANCED RASTER-TO-VECTOR BENCHMARK TESTING		5. TYPE OF REPORT & PERIOD COVERED Technical Report 2/86-6/86
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David D. Selden		8. CONTRACT OR GRANT NUMBER(s) DAA29-81-0100-1801
9. PERFORMING ORGANIZATION NAME AND ADDRESS BATTELLE COLUMBUS DIVISION 2030 M St., N.W. Washington, D.C. 20036		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060-5546		12. REPORT DATE June 1986
		13. NUMBER OF PAGES 69
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Defense Mapping Agency Hydrographic/Topographic Center Washington, D.C. 20415-0030		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution is unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Raster-to-Vector, Cartographic Data Capture, Benchmark Testing Scanning, Editing, Error Detection/Correction		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Battelle Columbus Division in cooperation with the U.S. Defense Mapping Agency and U.S. Army Engineer Topographic Labs has developed a methodology for testing and evaluation of automated cartographic editing functions provided by cartographic data capture systems. Editing functions include gap closure, spike removal, "snow" removal, node snapping, centerline adjustment, etc.. These functions are often implemented in three modes; automatic, computer-assisted and manual. The Battelle benchmark testing method focuses		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

on automated editing. The testing and evaluation method consists of a single analog (film positive) edit test sheet and a series of test procedures. The edit test sheet portrays a column of cartographic patterns with anomalous versions of these cartographic geometries arrayed in rows to their right. Different lineweights and pattern orientations are also utilized. The full range of editing functions are applied to the data to test for rates of success and sensitivity to geometry, lineweight and orientation.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

# PREFACE

This document was prepared under contract DAAG29-81-D-0100-1801 for the Defense Mapping Agency Hydrographic/Topographic Center, Washington, D.C., and the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia. The work was performed by Battelle's Columbus Laboratories, Washington, D.C., Operations. The Contracting Officer's Representative was Ms. Betty Mandel.



Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

## EXECUTIVE SUMMARY

Battelle Columbus Division (BCD) has completed a research and development project for the Defense Mapping Agency (DMA) under contract no. DAAG29 - 81 - D - 0100 - 1801 entitled "A Defense Mapping Agency (DMA) Advanced Raster-to-Vector Benchmark Testing." Three major tasks were successfully performed.

### 1) Development of Benchmark Testing Materials and Testing Methods

Battelle developed an Edit Test Sheet (ETS) and benchmark testing methodology for computer assisted and automatic error detection/correction capabilities in cartographic data capture systems.

- The Edit Test Sheet consists of a two dimensional matrix of geometric patterns which emulate basic cartographic symbologies resident on DMA maps. The left-most column of the ETS contains "perfect" geometric patterns and degraded (i.e., containing intentionally made imperfections) versions of each pattern are arrayed to the right.
- The benchmark testing methodology has five objectives:
  - to determine the success rates of computer assisted and automatic error detection/correction capabilities in cartographic data capture systems
  - to monitor system/process induced errors
  - to identify the types of errors which result from specific processing functions
  - to record times to completion for error detection/correction routines
  - to provide a means of comparison for alternative editing scenarios (e.g., manual interactive vs. computer assisted or automatic routines)



2) Implementation of the Edit Test Sheet (ETS) Benchmark Methodology

Battelle implemented the Edit Test Sheet (ETS) benchmark methodology on the Scitex Response 280 and Broomall Automated Graphic Digitizing System (AGDS) at the Defense Mapping Agency Hydrographic/Topographic Center. The main objective of this testing was to validate the ETS and benchmark testing methodology. The Scitex demonstrated a variety of error detection/correction routines. Success rates ranged from 67% (e.g., gap closure) to 100% (e.g., snow removal). Alternative editing scenarios were compared on the Scitex, demonstrating advantages of specific editing strategies (e.g., improved gap closure speed by automated means compared to interactive techniques). Testing of the AGDS highlighted this system's total lack of computer assisted or automatic error detection/correction capabilities.

3) Analysis of Benchmark Testing Results and the ETS Methodology

Battelle performed two analyses during the project. First, the benchmark testing results from the Scitex and Broomall data capture systems were analyzed in terms of individual performance and the two systems were compared. Second, the Edit Test Sheet (ETS) and benchmark testing methodology were analyzed.

- o The Scitex demonstrated a reasonable degree of success with available computer assisted and automatic error detection/correction software.

- The interactive editing capabilities of the Scitex are quite extensive in both raster and vector modes. The limited window of addressability in raster mode (approximately 1-2 square inches on a source document) is a weakness of the current approach. The vector editing system is somewhat awkward, particularly in comparison to many computer aided design (CAD) systems.
- Overall the Scitex shows promise in the area of automated error detection/correction. Fuller use of existing capabilities including the batch programming language will produce even better results.
- The Broomall AGDS has very limited computer assisted and automatic error detection/correction capabilities.
- The Broomall AGDS does have a comprehensive interactive editing capability in its Edit/Tag subsystem.
- In general, the Scitex is more advanced in the area of computer assisted and automatic error detection/correction than the Broomall AGDS. On the other hand, the vector interactive editing capability of the AGDS is superior to the interactive editing capabilities of the Scitex. Working together, these two systems could produce an overall improvement in data production capacity at DMA.

Analysis of the Edit Test Sheet (ETS) and benchmark testing methodology affirmed the validity of the approach. The basic objectives of the benchmark testing package were met:

- Its black and white image on clear film permits its application to manual digitizing, automatic line following and raster scanning data capture systems
- It contains representative samples of basic cartographic line symbologies

- Synthetically created geometric patterns in the ETS have consistent, specified lineweights and "error" dimensions. This facilitates controlled evaluation and analysis.
- A single test sheet minimizes the time required for test data input.
- Changes to the "perfect" geometric patterns are automatically recorded and can be compared with the original input.
- Repetitive testing of error detection/correction routines with modified tolerances provides quantitative measures success rates.

Battelle has several recommendations for further research and development based on the results of this project:

- Development of Enhanced Benchmark Testing Methods

Battelle believes that based on previous research, improvements and refinements to the benchmark testing approach are possible and will accrue real benefits to DMA. Battelle also recommends the creation of a digital version of the benchmark testing package and the design of an extension to the benchmark package to address the emerging technologies of pattern/symbol/text recognition in state-of-the-art cartographic data capture systems.

- Development of Enhanced Capabilities on the Scitex Response 280

Battelle recommends the development of enhanced data processing capabilities on the Scitex Response 280 at DMAHTC using available tools. Specifically, Battelle recommends that DMA interface an IBM PC to the Scitex Raster Edit workstation to support the following activities.

- interactive vector editing

- interactive vector digitizing and feature tagging
- production control and system administration
- on-line applications documentation
- on-line training sessions

Battelle recommends that batch processing/editing routines be developed on the Scitex using the batch programming language. These routines should be designed to implement a completely automated third shift for data production at DMA.

- Evaluation and Testing of Scitex and AGDS Applications Integration

Battelle recommends that DMA explore the feasibility of using the Scitex and AGDS data capture systems in an integrated, complimentary manner for data production. The effective use of each system for components of the data conversion process should be evaluated and tested. For example, the use of the Scitex for initial data entry and automated raster data editing, combined with the vector editing facilities of the Broomall AGDS might produce an overall improvement in productivity. These and other technology/system interfacing (using the Raster Reformatting software recently developed by DMA) could produce real benefits for DMA while using existing capabilities.

- Development of Automatic and Computer Assisted Editing Routines on the Broomall AGDS

Battelle recommends that DMA develop automatic or computer assisted routines to detect and correct data errors or anomalies on the Broomall AGDS. These capabilities will greatly improve the efficiency of data processing on the AGDS.

ACKNOWLEDGEMENTS

The author thanks the following individuals for their contributions: Trudee Izlar, Karen Krasner, and Stephanie Brown for secretarial and editing support; staff members of the Scitex Response 280 and Broomall AGDS at DMAHTC; Douglas Caldwell and Betty Mandel from the U.S. Army Engineer Topographic Laboratories who served as contract officers; and Art Noma, Rick Joy and Chung Hye Read from DMAHTC who made valuable suggestions and facilitated Battelle's work throughout the project.

## TABLE OF CONTENTS

	<u>Page</u>
1.0 INTRODUCTION.....	1
1.1 BACKGROUND.....	1
1.2 REVIEW OF EARLIER RESEARCH.....	1
1.2.1 A Defense Mapping Agency Raster-to-Vector Analysis.....	2
1.2.2 A Defense Mapping Agency Raster-to-Vector Benchmark Testing.....	2
1.3 PROBLEM DEFINITION.....	4
1.4 CURRENT RESEARCH OBJECTIVES.....	7
1.5 BENCHMARK TESTING ENVIRONMENT.....	8
1.5.1 The Scitex Response - 280.....	8
1.5.2 The Broomall Automated Graphic Digitizing System (AGDS).....	8
2.0 DEVELOPMENT OF BENCHMARK TESTING MATERIALS AND TESTING METHODS.....	9
2.1 DEVELOPMENT OF AN EDIT TEST SHEET.....	9
2.1.1 The Concept.....	9
2.1.2 The Content and Specification.....	10
2.1.3 Production Method.....	20
2.2 BENCHMARK TESTING METHODS AND PROCEDURES....	20
2.2.1 Benchmark Evaluation Objectives.....	20
2.2.2 Benchmark Test Procedures.....	22
2.2.3 Evaluation and Rating.....	23
3.0 BENCHMARK TESTING PACKAGE IMPLEMENTATION AND VALIDATION.....	24
3.1 SCITEX.....	24
3.1.1 Testing Procedures and Results.....	24
3.2 BROOMALL AGDS.....	28
3.2.1 Testing Procedures and Results.....	28

TABLE OF CONTENTS

	<u>Page</u>
4.0 ANALYSIS OF THE EDIT TEST SHEET (ETS) AND EVALUATION OF BENCHMARK TESTING METHODS.....	32
5.0 CONCLUSION.....	34
6.0 RECOMMENDATIONS FOR FURTHER RESEARCH AND DEVELOPMENT.....	35
REFERENCES	
APPENDIX A WORKFLOW DIAGRAM FOR BENCHMARK TESTING	
APPENDIX B BENCHMARK TESTING THE BROOMALL AGDS	
APPENDIX C ANALYSIS OF BENCHMARK TESTING RESULTS ON THE SCITEX RESPONSE 280 AND THE BROOMALL AGDS AT DMAHTC	
APPENDIX D A STATEMENT ON THE IMPORTANCE OF THE RESEARCH	

## 1.0 INTRODUCTION

Battelle Columbus Division (BCD) has completed a research and development project for the Defense Mapping Agency (DMA) under contract No. DAAG29-81-D-0100-1801 entitled "A Defense Mapping Agency (DMA) Advanced Raster-to-Vector Benchmark Testing". This technical report summarizes the research findings and prepares the groundwork for further investigations.

### 1.1 BACKGROUND

The Defense Mapping Agency is required to convert large quantities of paper maps or map compilations into a range of digital cartographic databases. This work is performed at DMA's two mapping centers in Brookmont, Maryland and St. Louis, Missouri employing, in general, a mix of commercially available technologies. Manual digitizing systems are currently applied with an increasing emphasis on raster scanning data capture techniques.\*

### 1.2 REVIEW OF EARLIER RESEARCH

In the past ten years the Defense Mapping Agency has supported research and development in cartographic data capture and attendant technologies. Work has focused on development of improved vectorization algorithms/software,<sup>1</sup> analysis of the analog-to-vector conversion process,<sup>2</sup> and development of benchmark testing methods for cartographic data capture systems.<sup>3</sup> This research has pursued improved comprehension of the cartographic data conversion process with particular emphasis on raster scanning technology and raster data processing. The goal remains to improve the efficiency of data production capabilities within the agency.

\*Therefore giving impetus to research of raster data processing and raster-to-vector conversion technology.



### 1.2.1 A Defense Mapping Agency Raster-to-Vector Analysis

Battelle Columbus Division initiated research activities in cartographic data capture at the Defense Mapping Agency by performing a comprehensive evaluation and analysis of the map conversion process at the agency's two mapping centers.<sup>2</sup> The results of this project include:

- Development of an analog-to-vector (A/V) conversion model (figure 1) which defines a framework for evaluation of cartographic data capture systems,
- A study of current and projected A/V conversion procedures/requirements at the DMA Hydrographic/Topographic Center and DMA Aerospace Center,
- A study of state-of-the-art automated cartographic data capture systems, and
- Development of a DMA A/V benchmark testing package and methodology.

### 1.2.2 A Defense Mapping Agency Raster-to-Vector Benchmark Testing

The focus of this second research effort<sup>3</sup> was to validate the benchmark materials and testing methods (developed during the previous research) on two data capture systems at DMA. A secondary emphasis was to evaluate the individual and comparative performance capabilities of these systems. The benchmark testing package was designed to analyze data conversion (i.e., scanning and vectorization) speed for basic cartographic inputs (e.g. contours), limited to "unsymbolized", or solid linear features. The impact of geometric and data volume characteristics on conversion performance was closely observed. The results of this project include:

# ANALOG-TO-VECTOR CONVERSION MODEL

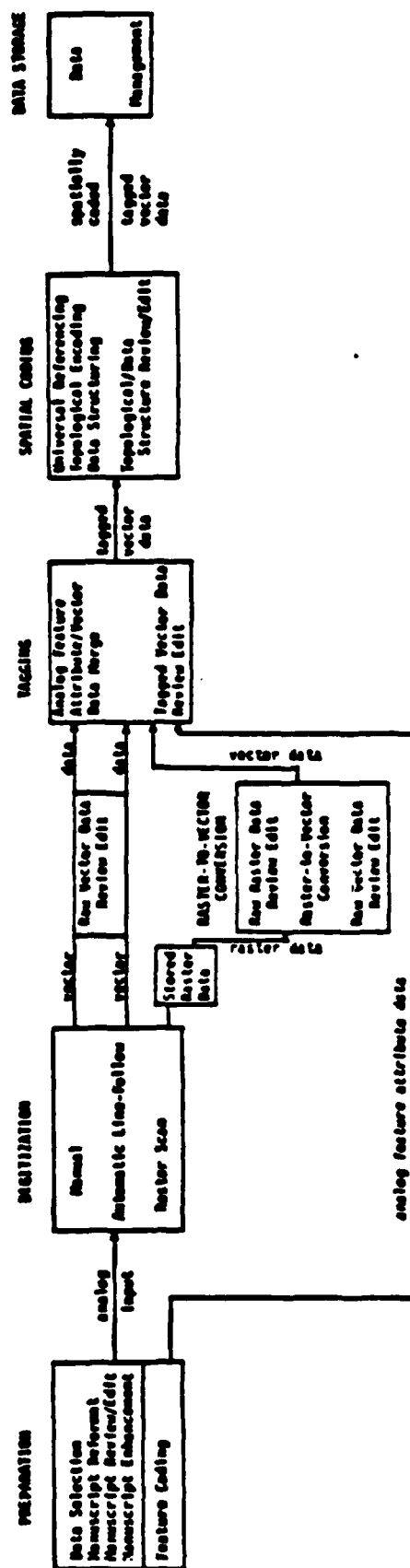


FIGURE 1.  
DMA ANALOG-TO-VECTOR CONVERSION MODEL

- A validation of the benchmark package and testing method,\*
- Test data demonstrating performance advantages of one data capture system over another at DMA,\*\* and
- A recommendation to produce and validate an analog-to-vector edit test sheet as an extension to the benchmark testing package.

These two research efforts provide a sound basis for the current project.

### 1.3 PROBLEM DEFINITION

Significant resources are required to detect and correct data errors during the creation of a digital cartographic database. The data capture process encounters data quality problems as a result of the characteristics of original analog source materials and as a byproduct of specific data capture techniques. Paper maps contain a wide range of characteristics which contribute to eventual digital data errors:

- breaks in lines,
- variable line image density,
- coalescing lines,
- miscellaneous dirt spots,
- variable line thickness, and
- overlapping or criss-crossing color areas/lines.

\*The research demonstrated the effectiveness of the benchmark testing methodology.

\*\*The Scitex Response 280 was demonstrated to perform raster-to-vector conversion faster than the Broomall AGDS at DMAHTC

Manual, automatic line following and raster scanning techniques perpetuate or exacerbate (to varying degrees) existing anomalies while contributing additional data quality problems:

- breaks in lines become digital coordinate gaps in "continuous" vectors (e.g., contours),
- variable line image density results in digital coordinate holes in raster data, eventually producing separate vectors where a single centerline is desired,
- line stubs or spikes become unintended coordinate data representing "non-existent" geographic features,
- coalescing lines become undefined "areas" or result in a single vector where multiple vectors are required (e.g., contours),
- miscellaneous dirt spots become unintended point, line or area feature data (typically resulting from raster scanning),
- variable line thickness results in "wavy" centerlines which may stray outside the bounds of specified positional accuracies, and
- thin line weights on analog sources occasionally result in missing data due to coarse raster scanning sampling resolution.

The identification and correction of data errors takes place throughout the A/V conversion process; source preparation, digitization, raster scanning, raster data editing, vector data editing, feature tagging and spatial structuring. Data capture systems provide a mixture of editing and quality assurance tools to address these problems. Three basic approaches to error detection/correction are:

- manual interactive computer graphic techniques-system operators scan through a computer display in a "frame-by-frame" sequence at an addressable resolution, visually identify errors, and make corrections interactively with a cursor, trackball or mouse,

- computer-assisted techniques - software routines identify data errors based on user-defined parameters (e.g., gaps .003" or less) and bring anomalies to the center of a CRT for interactive correction by a system operator, and
- automatic techniques - software routines identify and correct data errors based on user-specified parameters.

State-of-the-art cartographic data capture systems provide all three editing facilities with an emphasis on automatic processing and the utilization of human interaction where appropriate. However, even with modern facilities and the most advanced production software, significant amounts of human/time resources are required by all producers of digital cartographic data. This reflects the magnitude and complexity of digital cartographic files and the limited success rates of existing automatic data editing software.

The extent of the data error problem, combined with data producer's requirement for high production efficiencies and quality standards, are impetus to seriously evaluate the quality assurance capabilities of commercial data capture systems. This evaluation is pertinent to system acquisition decision-making or productivity enhancement planning for system developers or users. The Defense Mapping Agency "Advanced Raster-to-Vector Benchmark Testing" project has developed benchmark testing methods to assist the evaluation of cartographic data capture systems, emphasizing error detection/correction.

#### 1.4 CURRENT RESEARCH OBJECTIVES

Four research objectives were identified in the statement of work (SOW):

##### 1.4.1 Development of Benchmark Testing Materials and Testing Methods

The first objective was to develop materials and methods to test, evaluate and analyze the editing functions of cartographic data capture systems. The task included development of an Edit Test Sheet (ETS) from concept to content definition and specification. An ETS production method was also required.

##### 1.4.2 Benchmark Testing Implementation and Validation

The second objective was to implement the Edit Test Sheet on two raster data capture systems at the Defense Mapping Agency Hydrographic/Topographic Center (DMAHTC). The intent of this testing was to validate the ETS and benchmark testing methodology. A secondary goal was to evaluate individual system performance and do a comparative analysis of the two systems.

##### 1.4.3 Analysis of Benchmark Testing Results

The third objective was to analyze the qualitative and quantitative results of the benchmark testing. The analysis included the functionality of the ETS as a benchmark tool and the comparative individual editing performance of the two observed data capture systems. Evaluation of the benchmark testing methodology was also required.

#### **1.4.4 Recommendations for Further Research and Development**

The final objective of the project was to make recommendations for further research and development. Development of enhanced benchmark testing methods were to be addressed based on the research findings.

### **1.5 BENCHMARK TESTING ENVIRONMENT**

The benchmark testing validation was performed at DMAHTC on two raster data capture systems; the Scitex Response - 280 and the Broomall Automated Graphic Digitizing System (AGDS).

#### **1.5.1 The Scitex Response - 280**

The Scitex Response - 280 at DMAHTC is an integrated raster data capture system with raster and vector data processing capabilities. It has three major components: 1) a drum color raster scanner, 2) a color interactive computer graphics edit workstation and 3) a laser raster film plotter. DMAHTC maintains several edit workstations to support production of digital map products including hydrographic charts.

#### **1.5.2 The Broomall Automated Graphic Digitizing System (AGDS)**

The Broomall AGDS at DMAHTC is an integrated raster scanning and vector processing system consisting of three major subsystems: 1) a single color recognition flatbed raster scanner, 2) a vectorization subsystem and 3) an interactive (black and white) computer graphics edit/tag workstation. DMAHTC maintains several edit/tag workstations to support production of digital data products including Digital Terrain Elevation Data (DTED) and Digital Feature Analysis Data (DFAD).

## 2.0 DEVELOPMENT OF BENCHMARK TESTING MATERIALS AND TESTING METHODS

A primary objective of the research was to develop benchmark testing materials and testing methods to evaluate and analyze the error detection/correction capabilities of commercial cartographic data capture systems.

### 2.1 DEVELOPMENT OF AN EDIT TEST SHEET

The underlying concept of the edit testing methodology is that it correspond directly with the A/V conversion process. It has to measure error detection/correction capabilities commencing with capture of analog source materials and ending with production of centerline digital vector data.

#### 2.1.1 The Concept

The design of the analog Edit Test Sheet is based on the following precepts:

- provide an effective and flexible means of testing the three main types of data capture technology (i.e., manual, automatic line-following, and raster scanning) without requiring modifications to the material or hardware,
- represent "typical" cartographic source materials\* and facilitate controlled evaluation and analysis,
- limit the size and quantity to avoid the pitfalls of excessive time requirements for testing\*\*.

\*In this case DTED contour sheets, DFAD manuscripts, and hydrographic chart compilations were considered "typical" cartographic source material.

\*\*A common complaint heard from commercial system vendors is the extensive amount of time "required" by benchmark tests.



- represent the geometric characteristics typically found on analog cartographic source materials, and
- provide a means to track system/process induced errors and compute success rates of automatic and computer assisted editing functions.

### 2.1.2 The Content and Specification

A single Edit Test Sheet (figure 2) was created according to the principles stated above. It consists of a series of "synthetically" produced geometric patterns which emulate a range of elementary linear cartographic symbologies. The test sheet is plotted on a transparent film base with a series of "perfect" geometries in the left-most column and degraded (i.e., containing intentional errors and anomalies) versions of each geometric pattern arrayed to the right. The geometric patterns included in the test sheet from top to bottom are described in tables 1a, 1b, 1c, 1d, 1e, 1f, 1g, 1h.

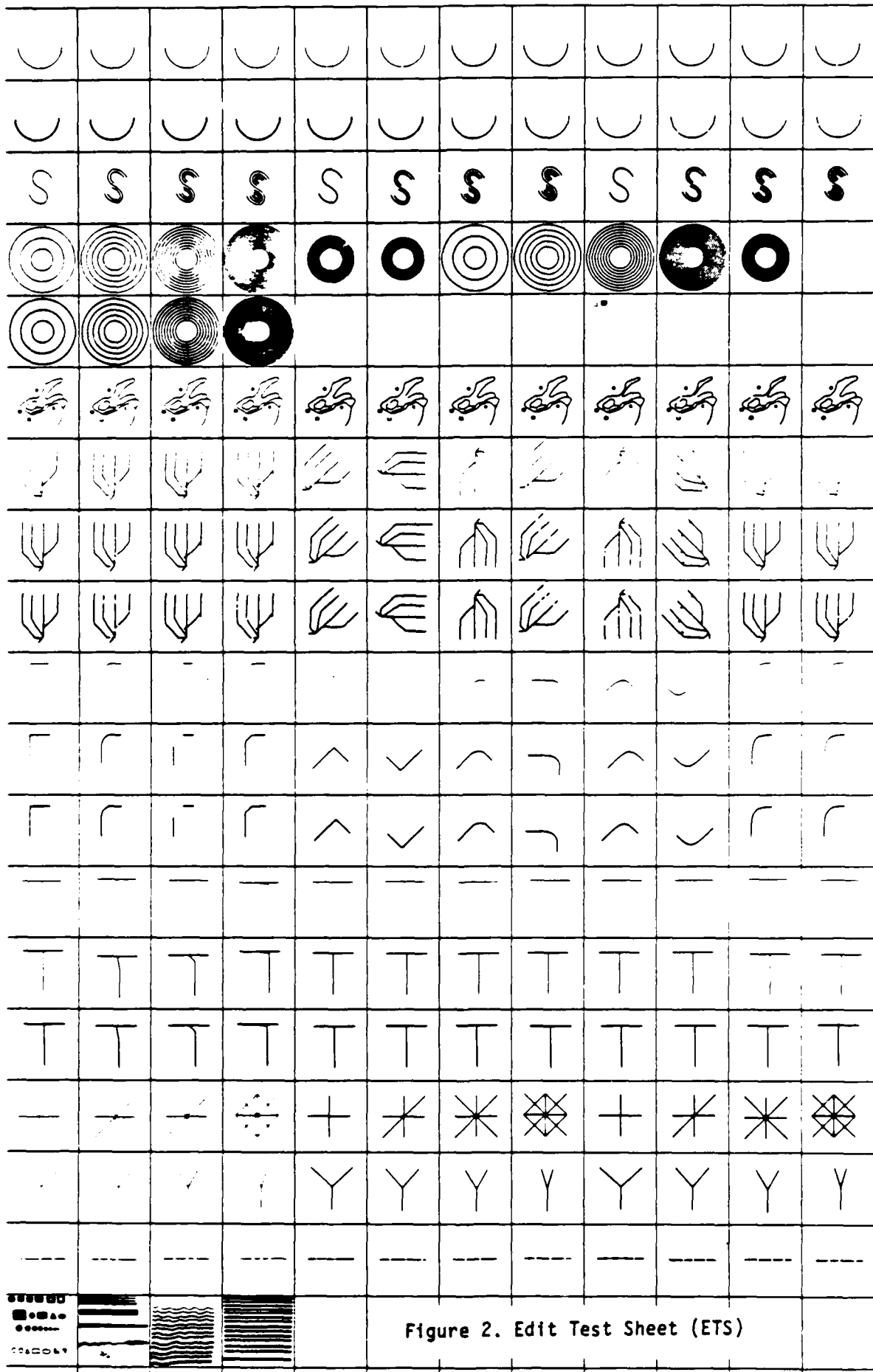


Figure 2. Edit Test Sheet (ETS)

ROW 1 Halfmoon Semi-circle; Represents section of cul- du-sac or portion of contour	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 1	.004"	N/A	N/A	N/A
COLUMN 2	.004"	GAP	.004"	N/A
COLUMN 3	.004"	GAP	.006"	N/A
COLUMN 4	.004"	GAP	.009"	N/A
COLUMN 5	.004"	GAP	.013"	N/A
COLUMN 6	.004"	GAP	.020"	N/A
COLUMN 7	.008"	N/A	N/A	N/A
COLUMN 8	.008"	GAP	.004"	N/A
COLUMN 9	.008"	GAP	.007"	N/A
COLUMN 10	.008"	GAP	.010"	N/A
COLUMN 11	.008"	GAP	.013"	N/A
COLUMN 12	.008"	GAP	.020"	N/A
ROW 2 COLUMN 1	.012"	N/A	N/A	N/A
COLUMN 2	.012"	GAP	.004"	N/A
COLUMN 3	.012"	GAP	.006"	N/A
COLUMN 4	.012"	GAP	.013"	N/A
COLUMN 5	.012"	GAP	.013"	N/A

ROW 2 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 6	.012"	GAP	.020"	N/A
COLUMN 7	.008"	GAPS	.005"	N/A
COLUMN 8	.008"	GAPS	.007"	N/A
COLUMN 9	.008"	GAPS	.014" .019"	N/A
COLUMN 10	.008"	GAPS	.021" .020"	N/A
COLUMN 11	.008"	GAPS	.003" .017"	N/A
COLUMN 12	.008"	GAPS	.018" .033"	N/A
ROW 3 S Curves; Represents parallel contour lines.				
COLUMN 1	.004"	COALESCENCE	N/A	N/A
COLUMN 2	.004"	COALESCENCE	N/A	.035"
COLUMN 3	.004"	COALESCENCE	N/A	.035"
COLUMN 4	.004"	COALESCENCE	N/A	.018"
COLUMN 5	.008"	COALESCENCE	N/A	N/A
COLUMN 6	.008"	COALESCENCE	N/A	.042"
COLUMN 7	.008"	COALESCENCE	N/A	.026"
COLUMN 8	.008"	COALESCENCE	N/A	.013"
COLUMN 9	.012"	COALESCENCE	N/A	N/A

ROW 3 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 10	.012"	COALESCENCE	N/A	.030"
COLUMN 11	.012"	COALESCENCE	N/A	.009"
COLUMN 12	.012"		N/A	.009"
ROW 4 Concentric Circles; Represents abstractions of of contours				
COLUMN 1	.004"	COALESCENCE	N/A	?
COLUMN 2	.004"	COALESCENCE	N/A	.093"
COLUMN 3	.004"	COALESCENCE	N/A	.045"
COLUMN 4	.004"	COALESCENCE	N/A	.026"
COLUMN 5	.004"	COALESCENCE	N/A	.012"
COLUMN 6	.004"	COALESCENCE	N/A	.002"
COLUMN 7	.008"	COALESCENCE	N/A	?
COLUMN 8	.008"	COALESCENCE	N/A	.091"
COLUMN 9	.008"	COALESCENCE	N/A	.042"
COLUMN 10	.008"	COALESCENCE	N/A	.021"
COLUMN 11	.008"	COALSCECE	N/A	.009"
ROW 5 Clumps of Dirt Spots				
COLUMN 1	.012"	COALESCENCE	N/A	?

ROW 5 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 2	.012"	COALESCENCE	N/A	.085"
COLUMN 3	.012"	COALESCENCE	N/A	.032"
COLUMN 4	.012"	COALESCENCE	N/A	.017 <sup>2</sup>
COLUMN 5	.001" .002" .003"	SNOW		N/A
COLUMN 6	.008" .010" .012"	SNOW		N/A
COLUMN 7	.014" .017" .019"	SNOW		N/A
COLUMN 8	.020" .022" .024"	SNOW		N/A
ROWS 6 Point, Line and Areal Feature Map; represents DFAD compilations	.004" .008" .012"	GAPS STUBS ORIENTATIONS		N/A
ROWS 7, 8, 9 Synthetic Drainage Patterns	.004" .008" .012"	GAPS STUBS ORIENTATIONS		N/A
ROWS 10, 11, 12 Corners represents any orthogonally joined line segments	.004" .008" .012"	ORIENTATIONS DEFORMATIONS		N/A
ROW 13 T Intersections; represent road intersections				
COLUMN 1	.004"	DEFORMATION	N/A	N/A
COLUMN 2	.004"	DEFORMATION	?	N/A
COLUMN 3	.004"	DEFORMATION	?	N/A
COLUMN 4	.004"	DEFORMATION	?	N/A
COLUMN 5	.004"	GAP	.005"	N/A
COLUMN 6	.004	GAP	.005"	N/A

ROW 13 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 7	.004"	GAP	.010"	N/A
COLUMN 8	.004"	GAP	.014"	N/A
COLUMN 9	.004"	STUB	.005"	N/A
COLUMN 10	.004"	STUB	.007"	N/A
COLUMN 11	.004"	STUB	.010"	N/A
COLUMN 12	.004"	STUB	.020"	N/A
ROW 14 COLUMN 1	.008"	DEFORMATION	N/A	N/A
COLUMN 2	.008"	DEFORMATION	N/A	
COLUMN 3	.008"	DEFORMATION	N/A	
COLUMN 4	.008"	DEFORMATION	N/A	
COLUMN 5	.008"	GAP	.005"	N/A
COLUMN 6	.008"	GAP	.009"	N/A
COLUMN 7	.008"	GAP	.010	N/A
COLUMN 8	.008"	GAP	.014	N/A
COLUMN 9	.008"	STUB	.005"	N/A
COLUMN 10	.008"	STUB	.007"	N/A
COLUMN 11	.008	STUB	.010"	N/A

ROW 14 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 12	.008"	STUB	.012"	N/A
ROW 15 COLUMN 1	.012"	DEFORMATION	N/A	?
COLUMN 2	.012"	DEFORMATION	N/A	?
COLUMN 3	.012"	DEFORMATION	N/A	?
COLUMN 4	.012"	DEFORMATION	.005"	N/A
COLUMN 5	.012"	GAP	.005"	N/A
COLUMN 6	.012	GAP	.007"	N/A
COLUMN 7	.012"	GAP	.010"	N/A
COLUMN 8	.012"	GAP	.014"	N/A
COLUMN 9	.012"	STUB	.005"	N/A
COLUMN 10	.012"	STUB	.007"	N/A
COLUMN 11	.012"	STUB	.009"	N/A
COLUMN 12	.012"	STUB	.019"	N/A
ROW 16 Crossing Intersections, Represents complex inter- section of line segments	.004" .008" .012"	N/A	N/A	N/A
ROW 17 Y Intersection; Represents merging of the streams				
COLUMN 1	.004"	ANGLE CHANGE	N/A	
COLUMN 2	.004"	ANGLE CHANGE	N/A	



ROW 17 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 3	.004"	ANGLE CHANGE	N/A	
COLUMN 4	.004"	ANGLE CHANGE	N/A	
COLUMN 5	.008"	ANGLE CHANGE	N/A	
COLUMN 6	.008"	ANGLE CHANGE	N/A	
COLUMN 7	.008"	ANGLE CHANGE	N/A	
COLUMN 8	.008"	ANGLE CHANGE	N/A	
COLUMN 9	.012"	ANGLE CHANGE	N/A	
COLUMN 10	.012"	ANGLE CHANGE	N/A	
COLUMN 11	.012"	ANGLE CHANGE	N/A	
COLUMN 12	.012"	ANGLE CHANGE	N/A	
ROW 18 Dashes; Represents simple dashing patterns				
COLUMN 1	.004	N/A	N/A	.017" .016" .016"
COLUMN 2	.004"	N/A	N/A	.021" .020" .020"
COLUMN 3	.004"	N/A	N/A	.025" .024" .024"
COLUMN 4	.004"	N/A	N/A	.028" .029" .029"
COLUMN 5	.008"	N/A	N/A	.020" .020" .020"
COLUMN 6	.008"	N/A	N/A	.020" .024" .025"

ROW 18 (continued)	Lineweight	Error Type	Error Dimension	Line Separation
COLUMN 7	.008"	N/A	N/A	.028" .029" .028"
COLUMN 8	.008"	N/A	N/A	.033" .033" .033"
COLUMN 9	.012"	N/A	N/A	.025" .025" .025"
COLUMN 10	.012"	N/A	N/A	.029" .029" .029"
COLUMN 11	.012"	N/A	N/A	.033" .032" .033"
COLUMN 12	.012"	N/A	N/A	.036" .037" .037"
ROW 19 Point and Line Symbols; Represents a range of basic <del>point and line symbols</del>				
COLUMN 1	.004" - ?	N/A	N/A	N/A
COLUMN 2	.002" - ?	VARIABLE LINE WIDTH	?	N/A
COLUMN 3	.001"-.030"	WAVY LINES	N/A	N/A
COLUMN 4	?	BROKEN LINES	?	N/A
COLUMN				
COLUMN				
COLUMN				
COLUMN				
COLUMN				
COLUMN				

### 2.1.3 Production Method

The Edit Test Sheet (ETS) was originally hand drawn in black ink on white paper. Specifically, all the "perfect" geometric patterns in the left hand column were hand drafted. The paper manuscript was optically scanned on the Scitex Super Scanner at DMAHTC. The digital raster image was interactively edited to remove imperfections in the geometric patterns and assure consistent dimensions. The original geometric patterns were converted to SYMPLACE\* files and replicated in the appropriate rows. Controlled errors (e.g., gaps) and anomalies were explicitly marked on the pattern duplicates using interactive computer graphic techniques. The final digital raster image was plotted on the Scitex laser plotter at a resolution of 39 points per millimeter (or approximately .001" accuracy). The Edit Test Sheet (ETS) is a film positive plot.

## 2.2 BENCHMARK TESTING METHODS AND PROCEDURES

Guidelines were developed to apply the Edit Test Sheet (ETS) for benchmarking the error detection/correction capabilities of cartographic data capture systems. Specific tests were developed and evaluation methods identified.

### 2.2.1 Benchmark Evaluation Objectives

The specific benchmark evaluation objectives are:

- to determine the success rates of computer-assisted error detection or automatic error detection/

\*SYMPLACE is a data format within the Scitex system which permits interactive placement of symbols on a working file from a symbol library.

correction routines for the characteristics presented on the Edit Test Sheet, including:

- gap identification
- gap identification and closure
- stub identification
- stub identification and removal
- snow\* identification
- snow identification and removal
- identification of non-orthogonal intersections or corners
- identification and correction of non-orthogonal intersections or corners
- identification of "hourglass" intersections
- identification and correction of "hourglass" intersections
- generation of the perimeter coordinates of solid point symbols
- successful vectorization of lines of any thickness or variable thickness
- production of centerline vectors
- clean vectorization of broken lines as a single vector,
- to monitor the system/process induced errors by observing changes or degradations to the "perfect" geometric patterns in the left hand column of the Edit Test Sheet,
- to identify the types of errors which result from specific processing functions (e.g., spikes resulting from vectorization),
- to quantify error detection/correction times for selected algorithms and symbologies, or for the entire ETS,

\*Miscellaneous dirt marks

- to provide a means of comparison for alternative editing scenarios (e.g., comparing speed and quality of manual frame-by-frame editing with computer-assisted or automatic processing).

### 2.2.2 Benchmark Test Procedures

Specific benchmark test procedures were developed for application of the Edit Test Sheet. Appendix A contains a workflow diagram of guidelines to benchmark test implementation. A summary of the benchmark test procedures applicable to scan data capture systems is presented below:

- a) scan digitize the Edit Test Sheet at .001" accuracy.
- b) perform a manual edit of a copy of the digital (raster) Edit Test Sheet and record the time to completion.
- c) "cutout" individual (or groups of similar rows) of geometric patterns and create separate digital (raster) files.
- d) run available error detection and error detection/correction routines against the individual raster files. Record successes and failures.
- e) vectorize the "original" digital raster Edit Test Sheet. Note deformities resulting from vectorization.
- f) vectorize individual geometric pattern files.
- g) perform a manual edit of a copy of the digital (vector) Edit Test Sheet and record the time to completion.
- h) run available error detection and error detection/correction routines against the individual vector files. Identify successes and failures.
- i) run alternative error detection and error detection/correction routines against raster and vector files. Record successes and failures and time requirements.

### 2.2.3 Evaluation and Rating

Evaluation of benchmark testing results is descriptive, graphic and quantitative. A written diary is compiled of the processing and specific edit tests applied to a data capture system. A graphic summary of the types of errors corrected or induced throughout the A/V conversion process is drawn on paper copies of the original Edit Test Sheet. Finally, timing statistics are compiled in tables and success rates are computed for computer assisted and automatic error detection/correction:

$$\text{Success rate} = \frac{\text{total number of errors detected/corrected}}{\text{total actual number of errors}}$$

For example: If the total number of gaps in row 1 is 15, and a gap closure function closes 10 gaps, the success rate for gap closure is 66%.

The written diary of data processing and benchmark testing provides a descriptive assessment of the overall capabilities of a cartographic data capture system in error detection/correction. The graphic summary depicts the types of errors induced by a system during data processing and indicates which errors were successfully identified and/or corrected. Finally, the success rates give a quantitative assessment of the capability of a given data capture system in identifying or correcting specific error types. It also provides a basis of comparison between systems.

### 3.0 BENCHMARK TESTING PACKAGE IMPLEMENTATION AND VALIDATION

The second objective of the research project was to test and validate the Edit Test Sheet (ETS). This was accomplished on the Scitex Response 280 and Broomall AGDS at the Defense Mapping Agency Hydrographic/Topographic Center (DMAHTC).

#### 3.1 SCITEX

The Edit Test Sheet (ETS) was run on the Scitex Response 280 at DMAHTC during a period of 50 hours in five individual sessions. This section of the report describes the test procedures, results, and any problems encountered.

##### 3.1.1 Testing Procedures and Results

Benchmark testing procedures developed earlier in the project were implemented, where appropriate, on the Scitex Response 280. The Edit Test Sheet (ETS) was scanned on the Scitex Super Scanner at a resolution of 39 points per millimeter, or approximately .001". Geometric patterns and anomalies were measured to confirm or modify the specifications of the analog ETS, and a series of processing/editing functions were performed.

Three forms of processing were performed on the Scitex Response 280: 1) ETS whole sheet processing, 2) ETS individual row editing, and 3) Alternative editing scenario testing. ETS whole sheet processing consisted of the combined interactive, computer assisted and automatic editing of the complete Edit Test Sheet. Interactive editing of the raw (i.e., directly off the scanner) raster ETS required approximately 5½ hours. An additional 4 hours were required to thin the raster image. Computer assisted

and automatic editing of the whole ETS was attempted and not completed. Interactive editing was required to complete the work because the computer-assisted techniques could not independently complete the processing requirement.

Individual rows of geometric patterns were "cut out" of an unedited version of the raw digital ETS and stored as separate files. Edit functions were tested against these individual files to facilitate quick results (compared to running edit functions against the entire ETS). First, two edit functions were tested against several individual raw raster data files:

- GAP, a computer assisted routine which locates gaps less than a specified tolerance, was run against rows 1 and 2 (halfmoon semi-circles), rows 13, 14, and 15 (T intersections), and row 18 (dashes).
- CLAR (Classified Retouch), an automated routine which identifies and deletes clusters of pixels within a specified min./max. tolerance, was run against row 5 (clumps of "dirt" spots).

Four editing functions were tested against individual thinned raster files:

- RSPIKE (Remove Spike), an automatic routine which identifies and deletes small pixel segments attached to longer pixel segments within specified tolerances, was run against rows 7, 8, and 9 (synthetic drainage patterns).
- GAP was run against rows 1 and 2 (halfmoon semi-circles), rows 13, 14, and 15 (T intersections), and row 18 (dashes).
- CONNECT, an automatic routine which closes gaps in pixel segments less than a specified tolerance, was run against rows 1 and 2 (halfmoon semi-circles), rows 13, 14, and 15 (T intersections), and row 18 (dashes).



- CLGAPDIG (Close Gap DIGIT), an automated routine which closes gaps in pixel segments less than a specified tolerance by creating links in DIGIT (i.e., vector x,y coordinate structure) file, was run against rows 1 and 2 (halfmoon semi-circles) and row 18 (dashes).\*

Alternative editing scenarios were tested by comparing different methods of identifying and correcting gaps in four individual rows of geometric patterns: 1) rows 1 and 2 (halfmoon semi-circles), 2) row 6 (point, line and areal maps), 3) rows 7, 8 and 9 (synthetic drainage patterns), and 4) row 18 (dashes). The following six editing methods were applied (except in the case of rows 7, 8 and 9 where only four methods were used):

- Raster Interactive Edit - frame by frame editing of a digital raster file using a crosshair cursor and function box commands
- Raster Edit Using GAP - the GAP function was run to identify gaps in the geometric patterns and bring them to the center of the CRT for interactive correction by an operator
- DIGIT Edit On a Raster File - the interactive placement of DIGIT (x,y) points at the beginning and terminating points of gaps in lines on a digital raster file. The execution of the DIGIT file onto the raster file completes the gap closure.
- VECTOR Edit On a Raster File - the interactive drawing of vector (x,y) lines in gaps on raster files. The vector links are converted to raster data to complete the gap closure.
- Automatic CONNECT Function - An automatic routine which closes gaps in thinned raster data within a specified tolerance.
- Automatic CLGAPDIG Function - An automatic routine which closes gaps in thinned raster data by creating links in a DIGIT file (x,y coordinate structure). The DIGIT file is converted to a raster format and superimposed on the original raster file to complete the gap closure.

\*Table 2 summarizes the results of the edit tests described above

EDIT FACILITY	ROWS			
	1 & 2 (raw raster)	18 (raw raster)	7, 8 & 9 (thinned raster)	18 (thinned raster)
	1 & 2 (thinned raster)			
GAP ID	97%	67%	-	67%
RSPIKE - (Spike Removal)	-	-	100%	-
CONNECT - (Gap Closure)	-	-	-	94%
CLGAPDIG - (Gap Closure)	-	-	-	100%
				93%

Table 2 Scitex Edit Testing Results

Table 3 describes the results of comparing alternative editing scenarios.

### 3.2 Broomall AGDS

The Edit Test Sheet (ETS) was run on the Broomall Automated Graphic Digitizing System (AGDS) at DMAHTC during a period of 32 hours in four individual sessions. This section of the report describes the test procedures, results, and any problems encountered.

#### 3.2.1 Testing Procedures and Results

The current version of software on the Broomall AGDS at DMAHTC has very limited error detection/correction capabilities, other than the interactive facilities offered on the edit/tag subsystem. To illustrate the limited automatic or computer assisted capabilities on the AGDS, the processing of the Edit Test Sheet (ETS) is described:

- The Edit Test Sheet (ETS) was scanned on the flatbed scanner at a resolution of .001", or 1000 lines per inch.
- The raw digital raster image of the ETS was reviewed for completeness and major anomalies. There were no error detection/correction capabilities in raster mode. Only visual inspection of the raster data image was permitted. The average linewidth in the raster file was measured. It was required as input to the vectorizer.
- Vectorization parameters were specified:
  - average width of lines
  - snow removal tolerance
  - hole (i.e., break within a raster line) closure tolerance

EDITING METHODS	ROWS			
	1 & 2	6	7, 8 & 9	18
Raster Interaction Edit	12 mins	30 mins	1 hr. 21 mins	17 mins
Raster Edit Using Gap	7 mins	5 mins	31 mins	8 mins
DIGIT Edit on Raster File	20 mins	30 mins	-	12 mins
digitizing executing DIGIT	15 mins 5 mins	23 mins 7 mins	- -	- -
VECTOR Edit on Raster File	25 mins	32 mins	-	14 mins
digitizing vectorizing DIGIT	15 mins 5 mins 5 mins	22 mins 5 mins 5 mins	- - -	- - -
Automatic CONNECT Function	3 mins	2.5 mins	10 mins	2 mins
run execute	2 mins 1 min	1 min 2 mins	- -	- -
Automatic CLGAPDIG Function	2 mins	3 mins	15 mins	4 mins
run execute	1 min 1 min	1 min 2 mins	- -	- -

Table 3. Results of Testing Alternative Editing  
Scenarios on the Scitex Response 280

- noise tolerance
- Vectorization was performed and produced a raw digital vector file.
- INODES was run. This function removed all unnecessary points along straight lines.
- THN was run. This function has three optional parameters:
  - 1) High Power Noise Filter - eliminates digital data noise defined by large directional change in a line over a small distance (sometimes referred to as spikes or loops).
  - 2) Low Power Noise Filter - eliminates digital noise defined by slight distance variation of a coordinate point from the straight line direction (otherwise resulting in "wavy" lines). A two pass filter is available to provide a measure of control over the degree of filtering.
  - 3) Stick Removal - eliminates two coordinate node segments which are not attached to any other segment in the data file.
- INODES was run a final time to remove any superfluous points remaining in the data file.
- The vector data file was transferred to the Edit/Tag subsystem for interactive review and editing. Standard interactive computer graphic techniques were applied to identify and correct data anomalies and errors. Error detection/correction of the ETS required 2 hours and 13 minutes.

Other than the parameters provided by the THN function, no error detection/correction functions were identified. The Broomall AGDS does not provide automatic or computer assisted capabilities for identifying or correcting anomalies or errors. This greatly limited the need for comprehensive testing using the Edit Test Sheet (ETS).

The remainder of the work on the AGDS tested the impact of modifying the THN parameters (i.e., High Power Noise Filter, Low Power Noise Filter, and Stick Removal) on the quality of the geometric patterns in the ETS. Appendix B describes the results.

#### 4.0 ANALYSIS OF THE EDIT TEST SHEET (ETS) AND EVALUATION OF BENCHMARK TESTING METHODS

The Edit Test Sheet (ETS) satisfies the basic objectives set forth in the design phase:

- Its black and white image (17" x 26") on clear film permits its application to manual digitizing, automatic line following and raster scanning data capture systems.
- It contains representative samples of basic cartographic line symbologies.
- Synthetically created geometric patterns in the ETS have consistent, specified lineweights and "error" dimensions. This facilitates controlled evaluation and analysis.

The benchmark testing method also satisfies the objectives set forth in the design phase:

- A single test sheet minimizes the time required for test data input.
- Changes to the "perfect" geometric patterns are automatically recorded and can be compared with the original input.
- Repetitive testing of error detection/correction routines with modified tolerances provides data needed to compute success rates.

Some limitations of the Edit Test Sheet (ETS) and benchmark testing methodology have been identified based on testing the approach on the Scitex and Broomall systems at DMAHTC. One limitation is the lack of absolute consistency between edit test data captured on different systems from the equivalent analog test sheet. All digitizing systems (including scanning) produce artifacts as a result of inaccuracies, different sampling

resolutions, and other hardware/software limitations. Therefore, comparison of raster-to-vector processing on two commercial data capture systems, for example, is complicated by differences in their respective raw raster data files. Another limitation of this benchmarking approach reflects a common difficulty of benchmarking in general; the amount of time required to fully test and explore the range of error detection/correction functions within most data capture systems. This is not solely a problem of the benchmark testing approach. Further consideration must be given to the question of what type of testing and how much testing is sufficient. A weakness of the current Edit Test Sheet (ETS) may be found in the lack of correspondence between the synthetic geometric patterns and real map symbology. Geometric patterns may be required which possess both consistent, known specifications and stronger similarity to cartographic symbols on typical DMA maps. The section on recommendations for further research and development will discuss potential approaches to improving the limitations identified above.



## 5.0 CONCLUSION

Battelle Columbus Division has successfully achieved the objectives set forth in the statement of work. Battelle has developed a benchmark testing methodology and Edit Test Sheet (ETS) which provide data on the error detection/correction capabilities of cartographic data capture systems. Battelle implemented the benchmark testing approach on the Scitex Response 280 and Broomall AGDS at DMAHTC to validate its basic design. Battelle's analysis of the benchmark testing highlighted the strengths and weaknesses of the approach to provide new insights into effective benchmark methodologies. These insights form the basis of recommendations for further research and development at DMA.

Cartographic data capture systems all possess generic tools which are applied to the analog-to-vector conversion process. Buyers of new systems and users of systems in production both require a comprehensive understanding of how this technology can be most effectively used to satisfy their unique requirements. Quality assurance and error detection/correction is a critical component of the A/V conversion process. Available data capture systems support a wide range of error detection/correction capabilities. Buyers and users need objective tools to evaluate and/or improve these capabilities. The edit test methodology developed by Battelle is a sound, objective approach to testing and evaluating error detection/correction facilities in cartographic data capture systems.

## 6.0 RECOMMENDATIONS FOR FURTHER RESEARCH AND DEVELOPMENT

Battelle recommends further research and development in four major areas:

### 1) Development of Enhanced Benchmark Testing Methods

Battelle recommends the development of a digital benchmark testing package. The Digital Benchmark Testing Package (DBTP) would contain the basic elements of the analog-to-vector benchmark testing package and Edit Test Sheet developed in previous Battelle research projects. The DBTP would test the raster-to-vector conversion capabilities of data capture systems and be used in conjunction with the analog benchmark testing materials. The advantages of a digital benchmark testing package would be greater control over the test pattern specifications, benchmark consistency between data capture systems, and automated means of benchmark testing evaluation. The DBTP would be designed for flexible application based on a hierarchy of objectives and constraints (e.g., testing only specific edit functions within a 10 hour period of time).

Battelle also recommends the development of a benchmark testing methodology for automated feature/symbol/text recognition and extraction. State-of-the-art data capture systems are introducing advanced capabilities to process analog cartographic source materials and produce "intelligent" databases by automatic and computer assisted means. Artificial intelligence and image processing techniques are being

applied to the complex challenge of converting paper maps into digital cartographic databases. Battelle recommends that the current benchmark testing methodology be extended to address these emerging advanced technologies. Battelle believes that the format of the Edit Test Sheet (ETS) serves as a model for developing benchmark testing methods for these advanced capabilities.

2) Development of Enhanced Capabilities on the Scitex Response 280

Battelle recommends the development of enhanced data processing capabilities on the Scitex Response 280 at DMAHTC. Battelle recommends development of capabilities in three areas:

- Implementation of an IBM PC interface with the Scitex would support the following activities:
  - interactive vector digitizing
  - interactive vector editing
  - production control and system administration
  - on-line applications documentation
  - on-line training sessions

At the Scitex User Group meeting in the autumn of 1985, Scitex announced the development of a communications link between the Scitex edit station and an IBM PC. Battelle recommends that DMA investigate recent progress in the linkage of IBM PC's to the Scitex. The activities described above

represent real advances which may only be realized via a microcomputer externally linked to the Scitex.

- Batch programming to develop an automated production shift should be investigated and implemented. The batch programming capability on the Scitex edit station offers an opportunity to link data processing and editing routines as a continuous process. Proper design could lead to a fully automated shift. Batch processing would occur overnight. Operators arriving for a morning shift would access processed files requiring only interactive editing. This would optimize the manual labor activities and promote greater automation. Examples of batch processing exist in a number of facilities operating Scitex systems. The U.S. Geological Survey, for example, uses batch processing extensively in the production of 1:100,000 - scale data files. Development of batch processing routines at DMA could be modeled after the procedures implemented at USGS or other companies.
- Testing and evaluation of all Scitex commands to determine their applicability to DMA data processing requirements is strongly recommended. The Scitex contains a wide range of processing and editing commands developed for various applications. Only a subset of these commands are currently in use at DMA. Comprehensive testing and applications evaluation is recommended. All tools currently provided by the Scitex which can be effectively applied to DMA requirements should be identified and used accordingly.

### 3) Evaluation and Testing of Scitex and AGDS Applications Integration

Benchmark testing of the Scitex and AGDS at DMAHTC reveals the strengths and weaknesses of both systems. In some ways these systems are complementary. For example, the Scitex has strong raster editing

capabilities including computer assisted and automatic functions. The AGDS has no raster editing or processing capabilities other than viewing functions. Otherwise, the AGDS has extensive interactive vector editing facilities. The Scitex's vector editing capabilities are somewhat awkward and limited. For a given application, if both systems were used where their capabilities are superior, the end result might be significantly improved throughput. Battelle recommends that this hypothesis be investigated and tested.

The Raster Reformatting Software recently developed by DMA offers a ready vehicle for evaluation of the potential for applications integration by the Scitex and AGDS. Optimal use can be explored by passing data in various stages of editing between the two systems. For example, a DTED compilation could be scanned on the Scitex Super Scanner and run through a batch routine of automated editing functions. Data could then be passed to the Edit Tag substation on the AGDS for interactive editing and feature tagging.

#### **4) Development of Automated and Computer Assisted Editing Routines on the Broomall AGDS**

Battelle recommends that DMA develop automatic or computer assisted routines to detect and correct data errors or anomalies on the Broomall AGDS. Many algorithms already exist and only require coding and implementation. Other algorithms need to be developed. Together, these routines would greatly improve the

efficiency of data processing on the AGDS. Limitation of tedious, time consuming manual interactive editing is a desirable and achievable goal. It can be accomplished with available techniques and technologies within the AGDS environment or in parallel with it.

## REFERENCES

- 1 Gillotte, M. J., "Raster Processing Techniques", Final Technical Report from Rome Research Corporation to Rome Air Development Center (February, 1979)  
  
Crane, Robert K., "Mini Raster-to-Vector Conversion", Final Report from Environmental Research and Technology, Inc. to U.S. Army Engineer Topographic Laboratories (September 1981) ETL-0269  
  
Mueller, C. S. M., and Christensen, A. H. J., "SCANLINE: A Computer Program for Converting Raster Maps into Line Maps", National Research Institute for Mathematical Sciences, Pretoria, South Africa, CSIR-TWISK-43, 1978
- 2 Selden, David D., Went, Burton, and Kleszczelski, Stan E., Final Technical Report of the Defense Mapping Agency (DMA) Raster-to-Vector Analysis, Battelle Columbus Laboratories, December 1984
- 3 Selden, David D., and Went, Burton, Final Technical Report of the Defense Mapping Agency (DMA) Raster-to-Vector Benchmark Testing, Battelle Columbus Laboratories, December 1984

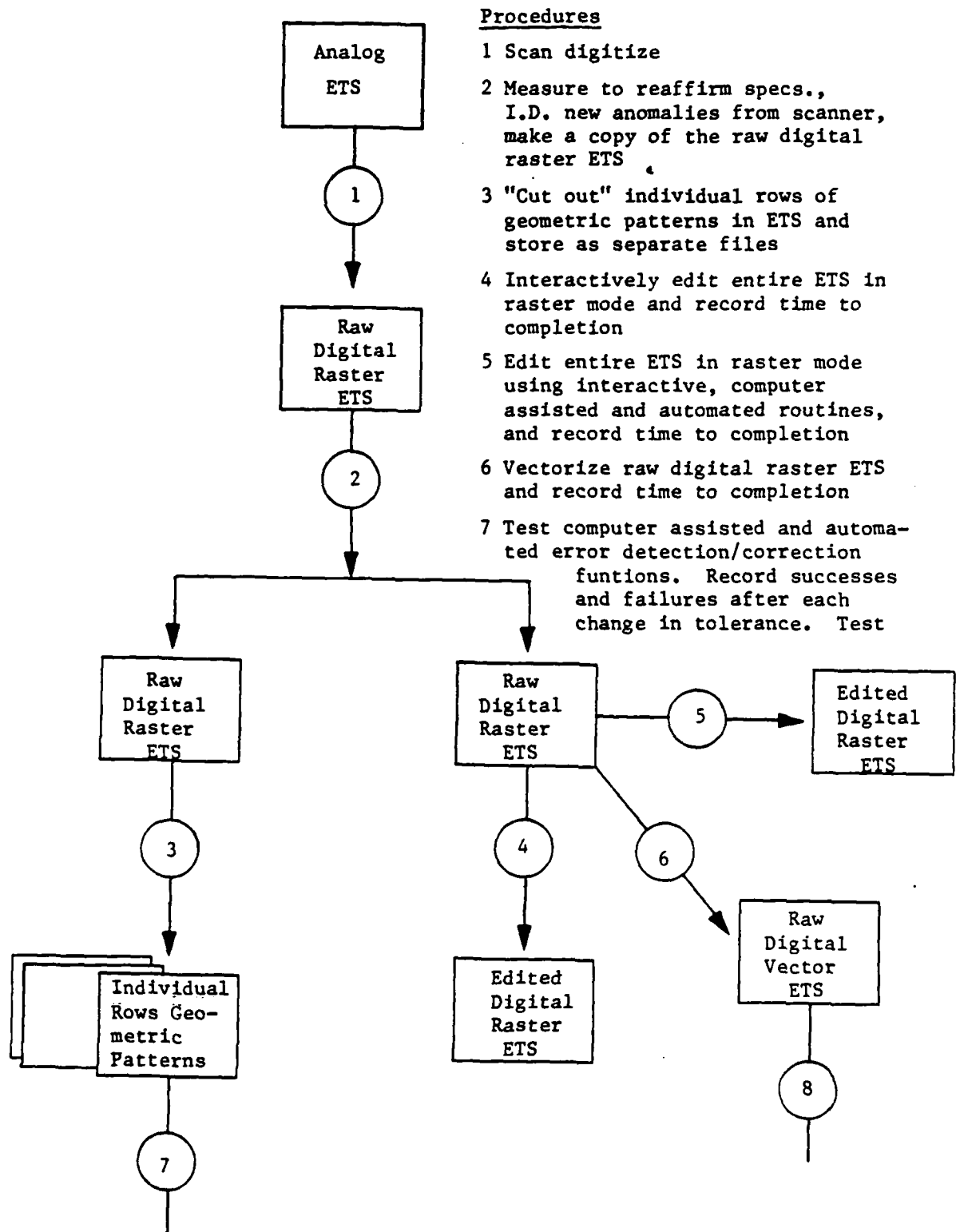
**APPENDIX A**

**WORKFLOW DIAGRAM FOR BENCHMARK**

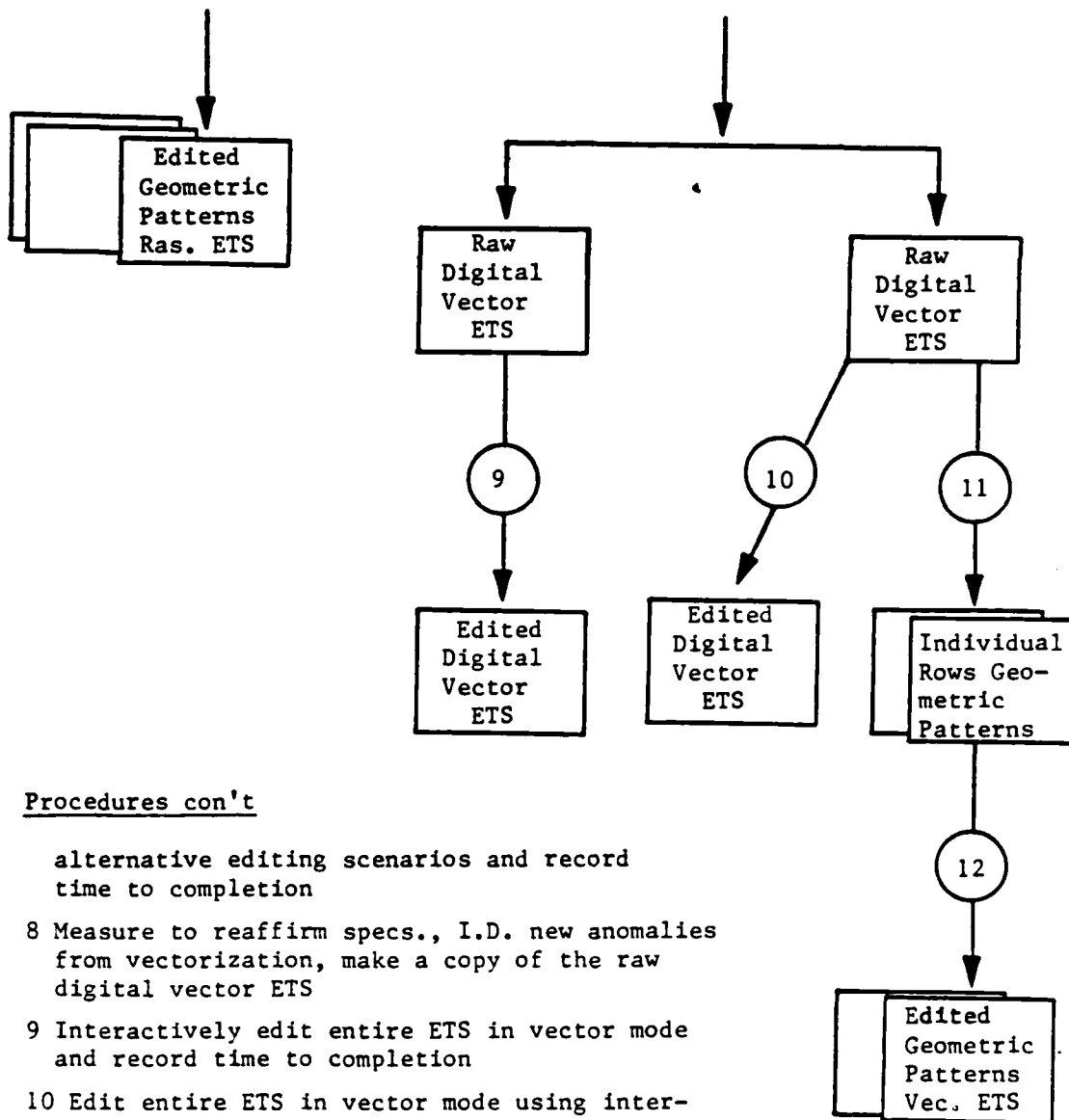
**TESTING (ETS)**



## WORKFLOW DIAGRAM OF THE EDIT TEST SHEET (ETS) BENCHMARK METHODOLOGY



"continued on following page"



#### Procedures con't

- alternative editing scenarios and record time to completion
- 8 Measure to reaffirm specs., I.D. new anomalies from vectorization, make a copy of the raw digital vector ETS
  - 9 Interactively edit entire ETS in vector mode and record time to completion
  - 10 Edit entire ETS in vector mode using interactive, computer assisted and automated routines, and record time to completion
  - 11 "Cut out" individual rows of geometric patterns in vector ETS and store as separate files
  - 12 Test computer assisted and automated error detection/correction functions. Record successes and failures after each change in tolerance. Test alternative editing scenarios and record time to completion

\* Note: The workflow diagram should be considered a guideline for for benchmark testing and not a rigorous step by step procedure. Testing will differ depending on the type of data capture technology and specific test objectives.

**APPENDIX B**

**BENCHMARK TESTING THE BROOMALL AGDS**

TESTING THN

Three geometric patterns from the Edit Test Sheet were run through the THN program on the AGDS using different filtering parameters. Rows 1 and 2 (halfmoon semi-circles), row 6 (point, line and areal maps), and row 16 (networks) were individually run through THN with different parameter settings and graphic copies of the resulting data were produced to illustrate the effect on geometric quality.

Rows 1 and 2 (halfmoon semi-circles) were first tested. Ten separate passes through THN were observed:

- 1) Raw Raster - direct, unaltered output from the scanner
- 2) Raw Vector/Pre-INODES - direct, unaltered output from the vectorizer
- 3) Raw Vector/Pre-INODES - direct, unaltered output from the vectorizer with nodes displayed
- 4) Vector/Post INODES - all remaining nodes on vector after INODES ran

The remaining passes represent post INODES testing.

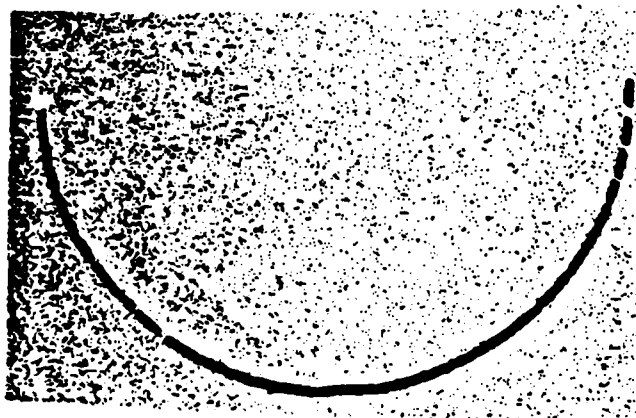
- 5) Vector THN/Sticks .024 - only stick removal function applied in THN
- 6) Vector THN/High Power .024 - only High Power Noise Filter applied in THN
- 7) Vector THN/Low Power (1) .016 - only Low Power Noise Filter with first pass applied in THN
- 8) Vector THN/Low Power (2) .016 - only Low Power Noise Filter with second pass applied in THN
- 9) Vector THN/Low Power (1) .016 (2) .004 - Low Power Noise Filter with pass (1) and (2) applied in THN
- 10) Vector THN/High Power .016/Low Power (1) .008 (2).002 - High Power Noise Filter and Low Power Noise Filter with pass (1) and (2) applied in THN

Note that the numbers following the filter establish the tolerance in thousandths of an inch. Figures 3 and 4 are screen copies of the geometric pattern in row 2 column 11 (R2C11) of the ETS subsequent to each pass prior to THN or after THN.

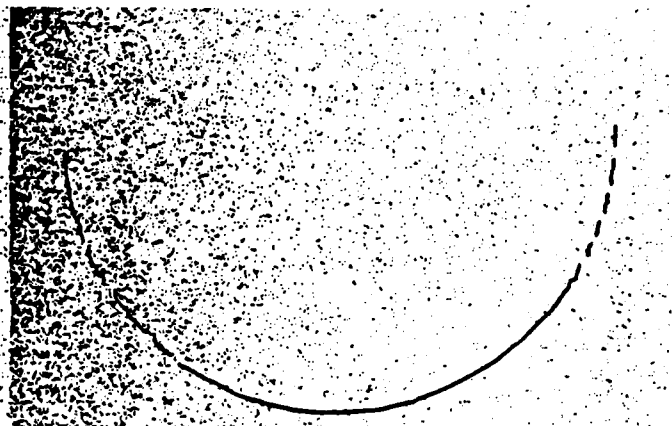
Row 6 (point, line and areal maps) was tested. Thirteen separate passes prior to THN, or after modified THN runs, were observed:

- 1) Raw Raster
- 2) Raw Vector/Pre-INODES
- 3) Raw Vector/Pre-INODES Displayed
- 4) Vector/Post INODES Displayed
- 5) Vector THN/High Power .005
- 6) Vector THN/High Power .024
- 7) Vector THN/Low Power (1) .004
- 8) Vector THN/LOW Power (1) .016
- 9) Vector THN/Low Power (2) .016
- 10) Vector THN/Low Power (1) .016 (2) .004
- 11) Vector THN/Low Power (1) .004 (2) .016
- 12) Vector THN/High Power .016/Low Power (1) .008 (2) .002
- 13) Vector THN/High Power .020/Low Power (1) .005  
(2) .0035

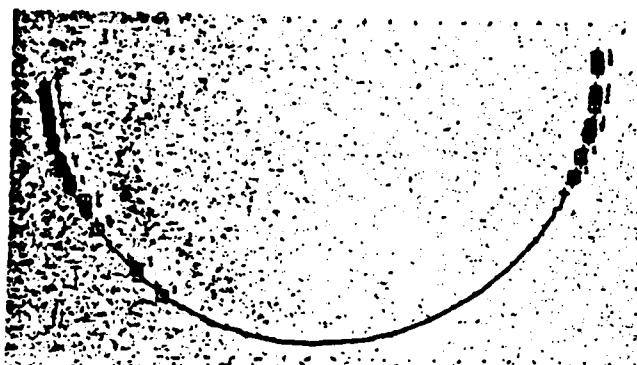
Figures 5, 6, 7, and 8 are copies of the geometric pattern in row 6, column 11 (R6C11) of the ETS subsequent to each pass prior to THN or after modified THN runs.



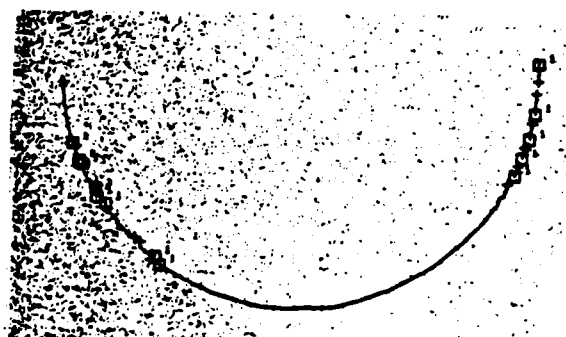
R2C11 Raw Raster



R2C11 Raw Vector/Pre-INODES



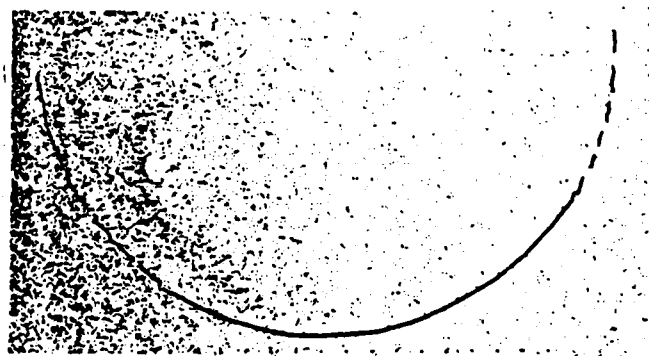
R2C11 Raw Vector/Pre-INODES  
(nodes displayed)



R2C11 Vector/Post INODES  
(nodes displayed)



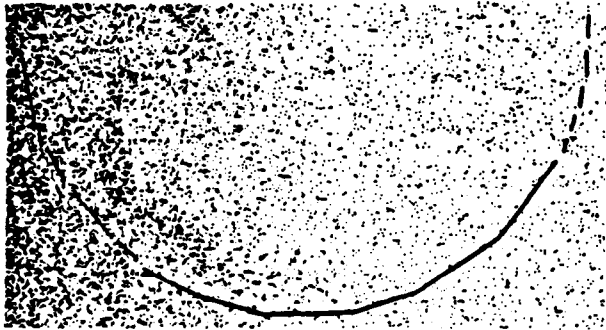
R2C11 Vector THN/Sticks .024



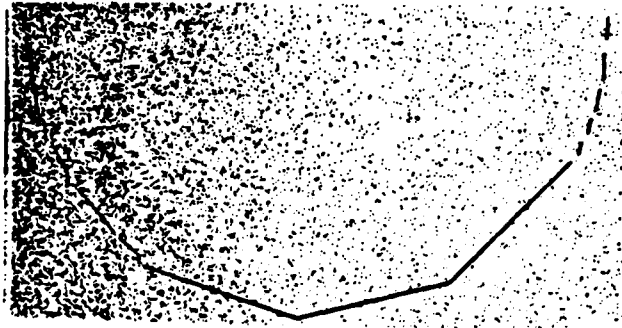
R2C11 Vector THN/High Power .024

FIGURE 3

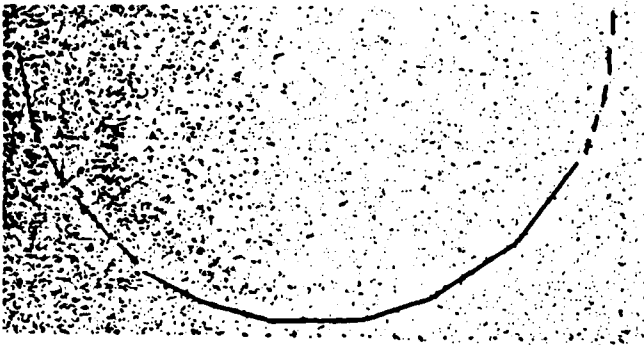
Testing THN - Rows 1 and 2 (halfmoon semi-circles)



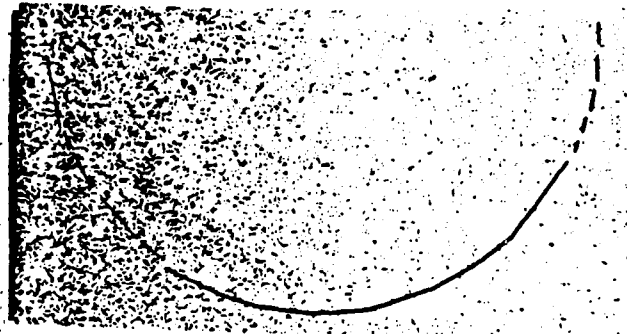
R2C11 Vector THN/Low Power (1) .016



R2C11 Vector THN/Low Power (2) .016



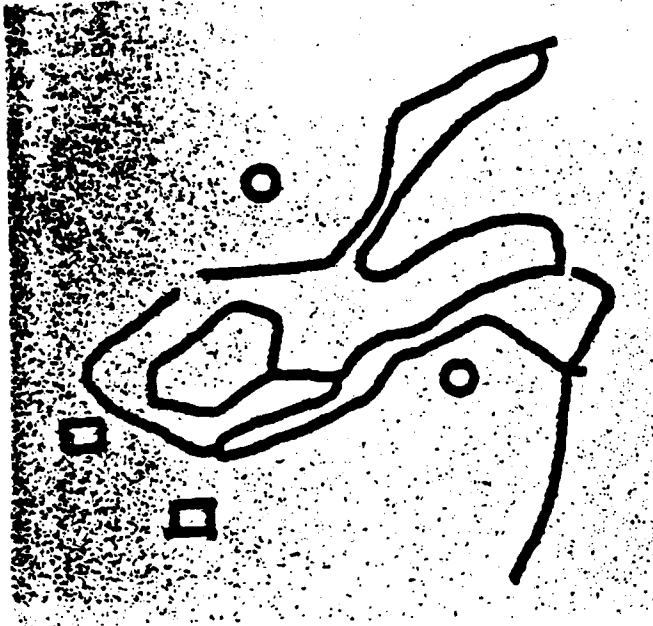
R2C11 Vector THN/Low Power (1) .016  
(2) .004



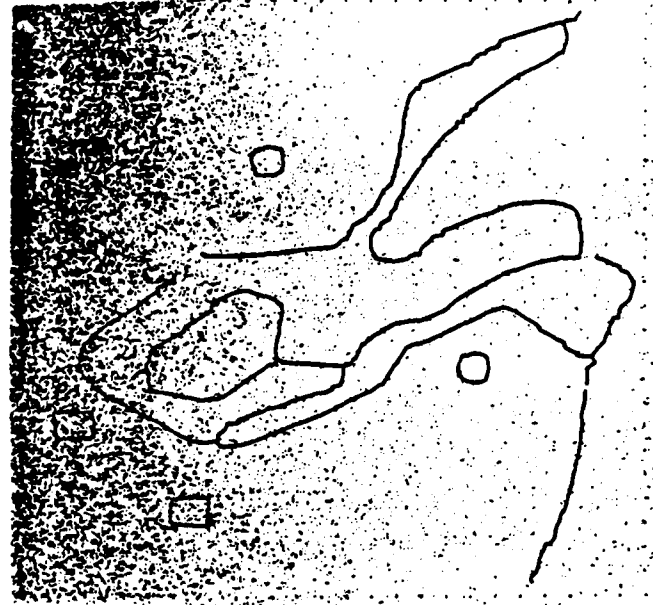
R2C11 Vector THN/High Power .016  
/Low Power (1) .008  
(2) .002

FIGURE 4

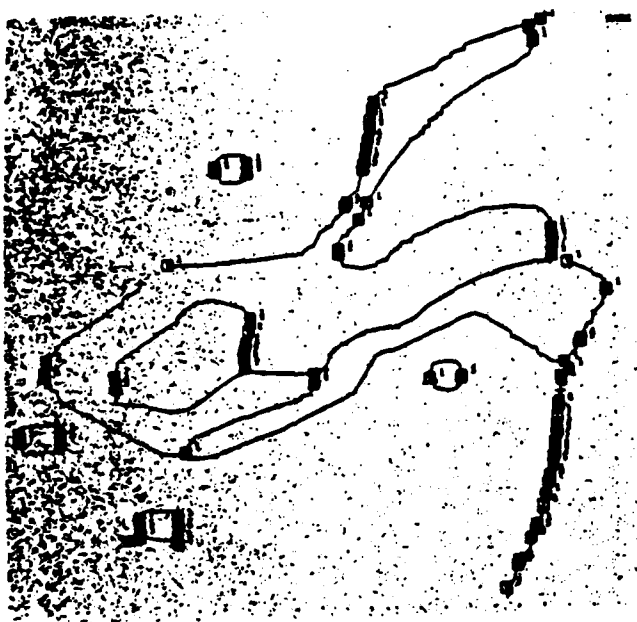
Testing THN - Rows 1 and 2 (halfmoon semi-circles)



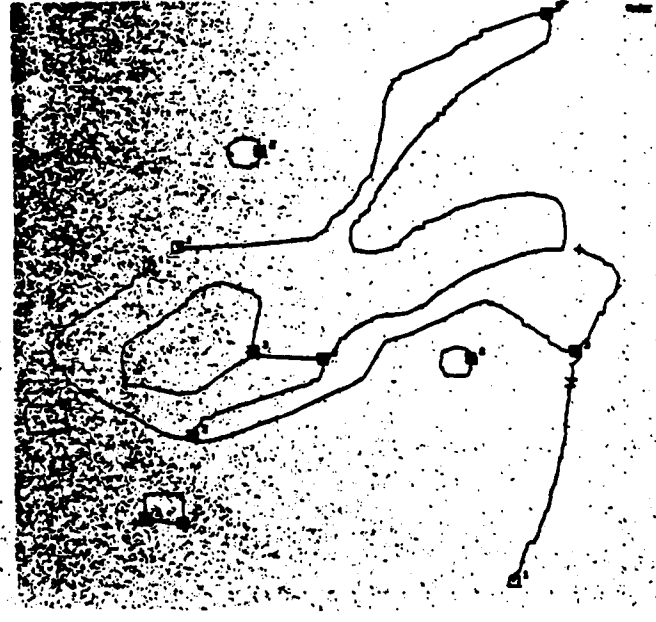
R6C11 Raw Raster



R6C11 Raw Vector/Pre-INODES



R6C11 Raw Vector/Pre-INODES  
(nodes displayed)

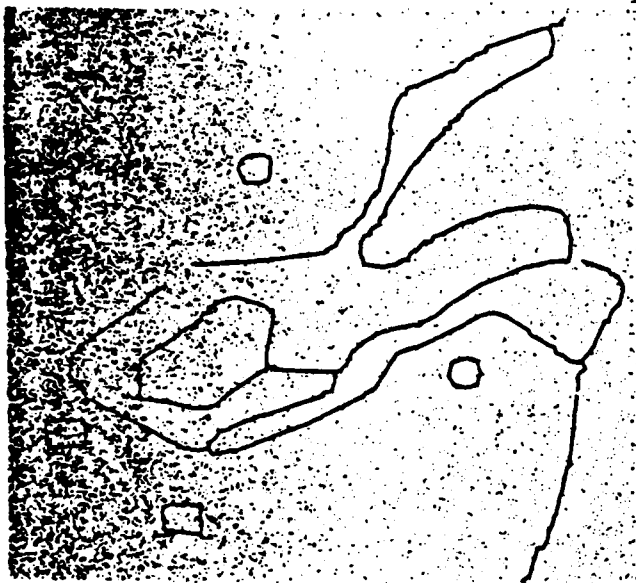


R6C11 Vector/Post INODES  
(nodes displayed)

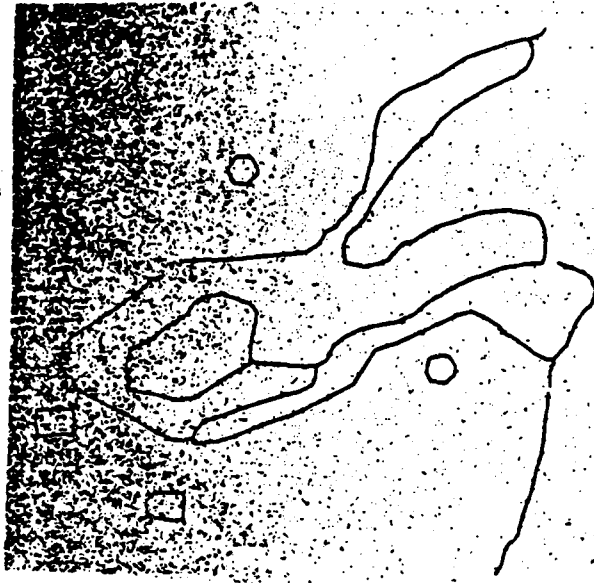
FIGURE 5

Testing THN - Row 6 (point, line and areal maps)

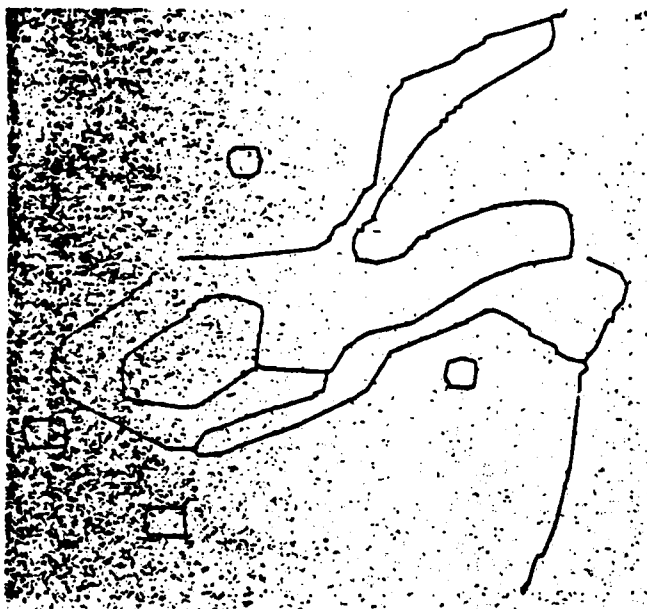




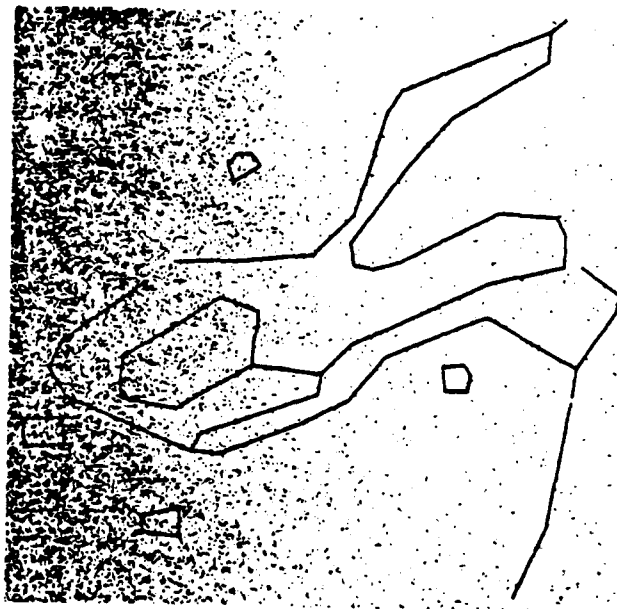
R6C11 Vector THN/High Power .005



R6C11 Vector THN/High Power .024



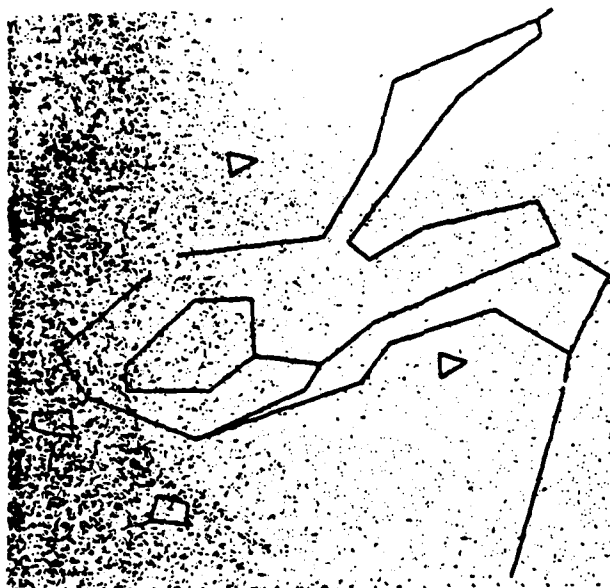
R6C11 Vector THN/Low Power (1) .004



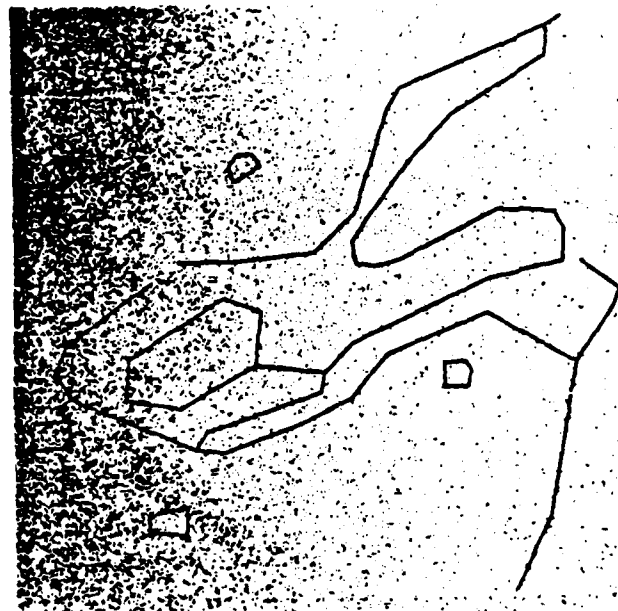
R6C11 Vector THN/Low Power (1) .016

FIGURE 6

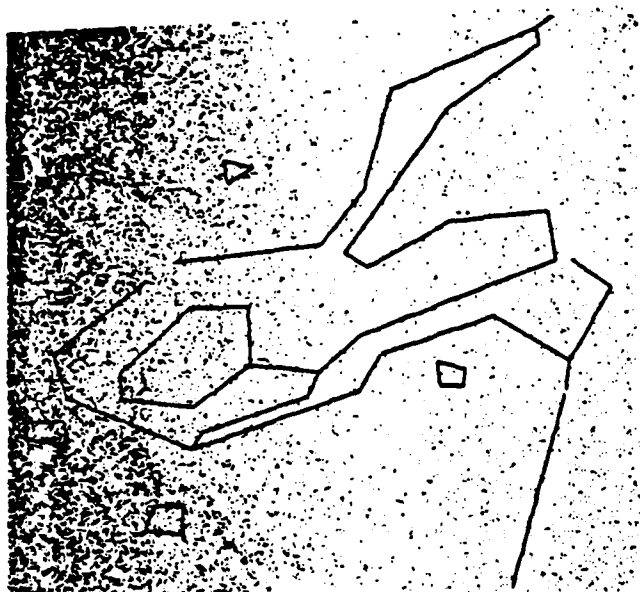
Testing THN - Row 6 (point, line and areal maps)



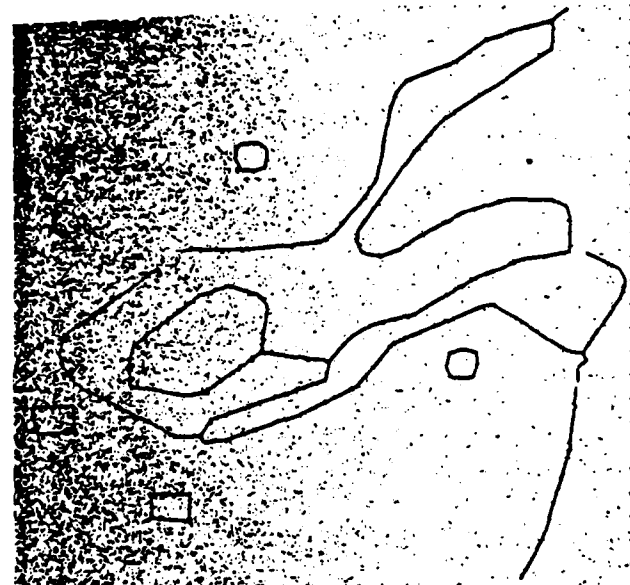
R6C11 Vector THN/Low Power (2) .016



R6C11 Vector THN/Low Power (1) .016  
(2) .004



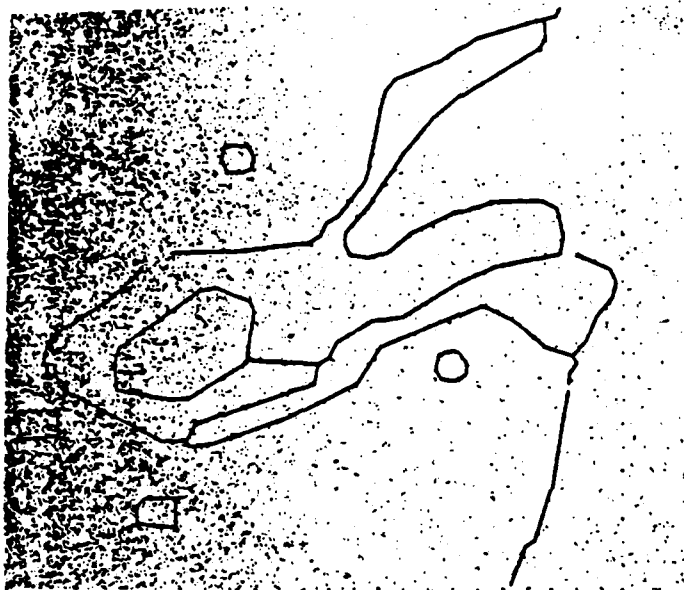
R6C11 Vector THN/Low Power (1) .004  
(2) .016



R6C11 Vector THN/High Power .016  
/Low Power (1) .008  
(2) .002

FIGURE 7

Testing THN - Row 6 (point, line and areal maps)



R6C11 Vector THN/High Power .020  
/Low Power (1) .005  
(2) .0035

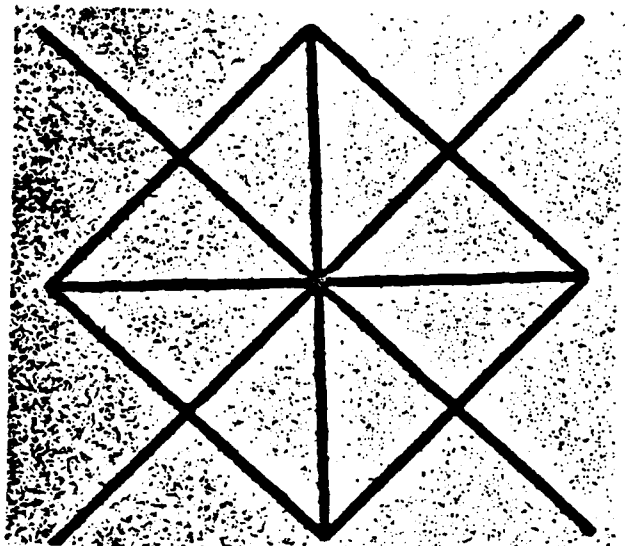
FIGURE 8

Testing THN - Row 6 (point, line and areal maps)

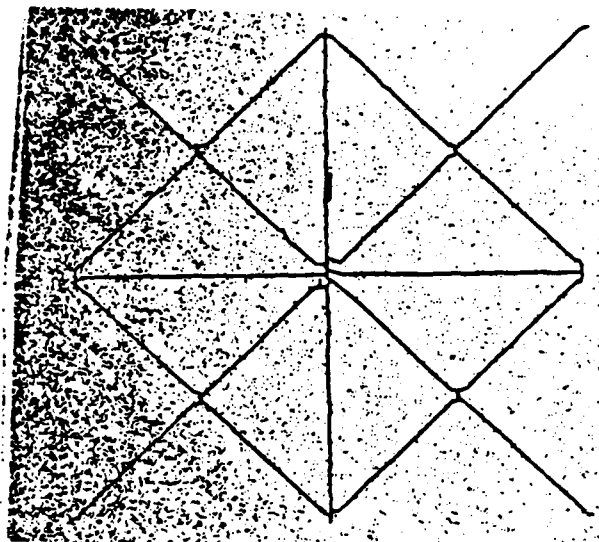
Row 16 (networks) was tested. Six separate passes were observed prior to THN or after modified THN runs:

- 1) Raw Raster
- 2) Raw Vector
- 3) Vector THN/Low Power (1) .016
- 4) Vector THN/Low Power (2) .016
- 5) Vector THN/High Power .020/Low Power (1) .005 (2) .0035
- 6) Vector THN/High Power .016/Low Power (1) .008 (2) .004

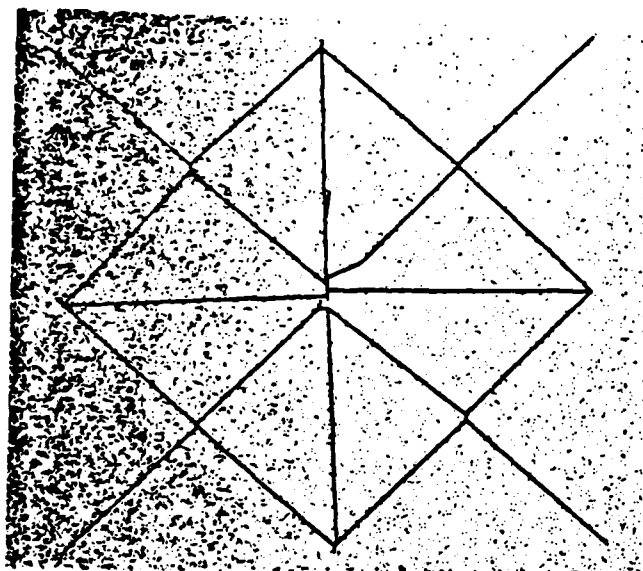
Figures 9 and 10 are copies of the geometric pattern in row 16, column 12 (R16C12) of the ETS subsequent to each pass prior to THN or after modified THN runs.



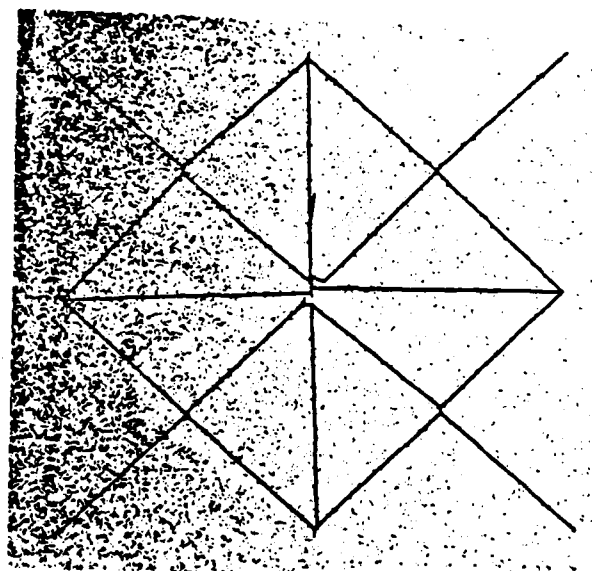
R16C12 Raw Raster



R16C12 Raw Vector

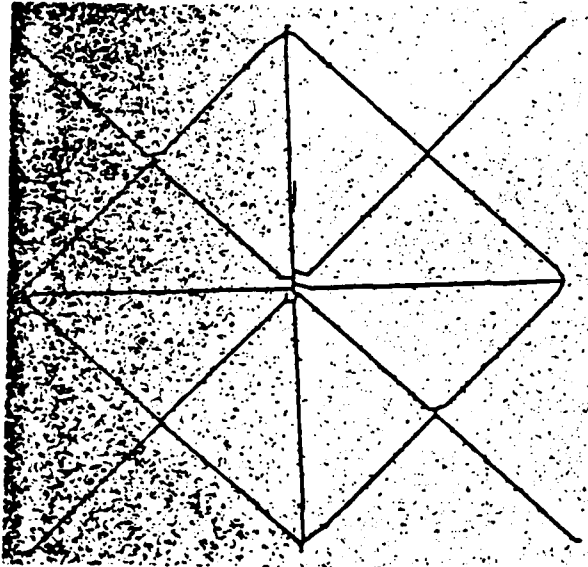


R16C12 Vector THN/Low Power (1) .016

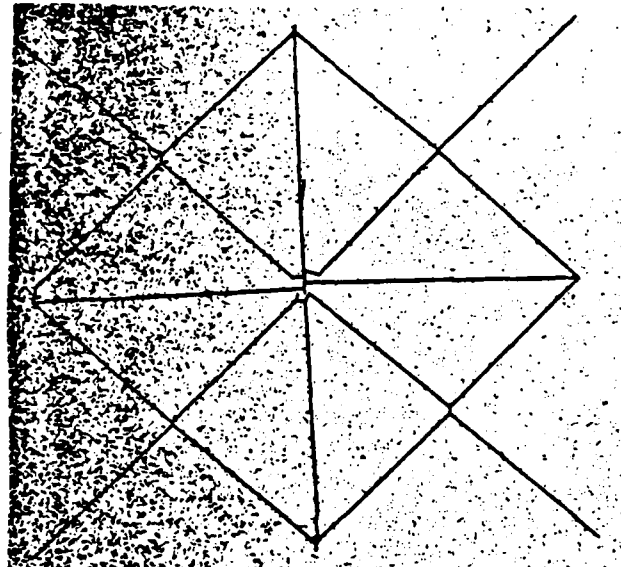


R16C12 Vector THN/Low Power (2) .016

FIGURE 9  
Testing THN - Row 16 (networks)



R16C12 Vector THN/High Power .020  
/Low Power (1) .005  
(2) .0035



R16C12 Vector THN/High Power .016  
/Low Power (1) .008  
(2) .004

FIGURE 10

Testing THN - Row 16 (networks)

**APPENDIX C**

**ANALYSIS OF BENCHMARK TESTING RESULTS  
ON THE SCITEX RESPONSE 280  
AND THE BROOMALL AGDS AT DMAHTC**

## ANALYSIS OF BENCHMARK TESTING RESULTS

The benchmark testing of the Scitex Response 280 and Broomall AGDS using the Edit Test Sheet (ETS) produced results which characterize their error detection/correction capabilities. This section analyzes the error detection/correction capabilities of each system and compares them.

### ANALYSIS OF BENCHMARK TESTING RESULTS - SCITEX

The Scitex Response 280 has a range of error detection/correction capabilities which work in interactive, computer assisted and automatic modes. Several of these functions were tested with the Edit Test Sheet (ETS) (e.g., GAP, RSPIKE, CONNECT, CLGAPDIG, CLAR). Each edit function had a success rate between 67% and 100%. Several difficulties were identified during the testing. Vertical gaps are neither identified nor corrected by the GAP, CONNECT or CLGAPDIG functions. Most functions fail to identify or correct errors whose size is equivalent to a specified tolerance (e.g., GAP fails to identify gaps .004" wide if the GAP tolerance is specified at .004"). No functions were capable of identifying or correcting non-orthogonal intersections and "hourglass" intersections. The Scitex also failed to generate the perimeter coordinates of solid point symbols and had difficulty producing the centerline of thick lines. It also failed to produce a single centerline vector from a dashed or broken line.

Most data errors and anomalies can be corrected on the Scitex interactively in raster or vector modes. Both modes have limitations. Perhaps the most severe limitation of the raster interactive editing mode is the limited window of addressability. Editing is limited to an area equivalent to approximately 1-2



square inches on a source document. Not only does this slow the manual editing process but it makes it difficult to determine where one is editing relative to the entire file. An edit index display is provided to assist in referencing the complete file but it is far less effective than many computer aided design (CAD) two screen "master/slave" facilities. The vector editing facility allows for interaction with individual points in a file. This offers certain complementary advantages over the raster editing. However, these interactive routines are somewhat awkward, particularly when compared to many CAD interactive vector editing workstations.

Overall the Scitex Response 280 shows promise in the area of computer assisted and automatic error detection/correction. Several basic functions currently exist and work reasonably well as described in section 3.1. A wider selection of functional automatic and computer assisted editing functions is needed. A considerable number of functions reside on the Scitex which are either nonfunctional or poorly understood. For example, the DIANA functions have been included in several versions of the Scitex software but have not been applied at DMA (as of the time of the benchmark testing in the autumn of 1985). Another problem is the availability of functions which appear to perform equivalent functions (e.g., MIDLINE, SKELTON, THN or CONNECT and CLGAPDIG). Confusion results from this apparent duplication. Part of the problem resides in the lack of applications oriented documentation (i.e., documentation which describes the application of Scitex functions for specific DMA requirements). Part of the problem is found in maintenance of older, less functional edit/processing routines which have been replaced with improved versions. Unfortunately, it is not always clear which function is current, better than another, or more appropriate for a given application. This is particularly true for new operators lacking experience with the system.

The Scitex Response 280 has the flexibility and basic structure to be improved in the area of automatic and computer assisted error detection/correction. Opportunities for improvement reside in the further development of capabilities by the company and in more effective use of available tools/routines by DMA. A discussion of opportunities for improvement is presented in the section of recommendations for further research and development.

## ANALYSIS OF BENCHMARK TESTING RESULTS - AGDS

The Broomall AGDS has very limited capabilities in computer assisted or automatic error detection/correction. Few error detection/correction capabilities are provided other than the automatic removal of miscellaneous two coordinate strings (i.e., Sticks) and line smoothing produced by the noise filters in the THN routine. However, a fairly complete system of interactive editing facilities is available on the Edit/Tag subsystem on the AGDS. Twin screen ("master/slave") editing is supported for flexible graphics manipulation. This includes the ability to create up to 32 "quepoints," or sub-editing windows per screen.

The parameters in the THN routine (i.e., High Power Noise Filter, Low Power Noise Filter and Stick Removal) directly affect the geometry of vector data. As indicated by the figures in section 3.2, THN can smooth out the "wrinkles" created by the vectorization process to produce acceptable linear representations. Alternatively, the smoothing can be too strong and produce exaggerated geometric configurations. THN does not, however, detect or correct most of the errors typically found in digital cartographic data (i.e., gaps, spikes, coalescence). These anomalies must be corrected by human operators using interactive computer graphic techniques.

Although the current version of software on the Broomall AGDS (at DMAHTC) does not provide capabilities for automatic or computer assisted error detection/correction, there are opportunities for improvement. A discussion of areas of potential improvement is presented in the section of recommendations for further research and development.

## A COMPARATIVE ASSESSMENT OF THE SCITEX AND AGDS

Benchmark testing and observation of the Scitex Response 280 and Broomall AGDS at DMAHTC provides ample evidence about their comparative capabilities in computer assisted or automatic error detection/correction. In their current state it is obvious that the Scitex is superior to the AGDS in the area of computer assisted or automatic error detection/correction. However, even though the Scitex has extensive interactive raster and vector editing tools, the AGDS appears to provide more effective capabilities in the area of interactive vector editing. Both systems induce errors or anomalies to digital data by raster scanning or vectorization. The use of color in the Scitex greatly assists the editing process and the lack of color adversely impacts editing efficiency on the AGDS. Both systems are relatively "user friendly" with text prompts on both systems, and the function box on the Scitex. The use of indirect command files (a DEC capability) on the AGDS and batch programming on the Scitex provide both systems with an efficient means to link editing/processing routines.

Both the Scitex and AGDS have strengths and weaknesses in the area of error detection/correction. There are opportunities for improving the performance of each system. There are also opportunities to improve overall performance by integrating the functions of both systems and taking advantage of inherent strengths in each system. A discussion of the opportunities for improving overall performance is presented in the section of recommendations for further research and development.

**APPENDIX D**

**A STATEMENT ON THE IMPORTANCE  
OF THE RESEARCH**

THE IMPORTANCE OF THE RESEARCH

Battelle believes that the development of benchmark testing methods and their implementation on data capture systems at DMA directly benefit the mapping agency in two important ways. First, it provides DMA with an objective methodology for evaluating the performance of commercial data capture systems for their unique requirements.

Second, it provides an opportunity to improve the productivity of data capture systems in use at DMA. The benchmark testing package facilitates experimentation with the data processing/editing tools available on these data capture systems. Experimentation is a necessary exercise to explore the range of capabilities and their appropriate use for DMA requirements. Only through experimentation, testing and evaluation of alternative data processing techniques, will optimal data capture methods be developed.

END

10-8%

DTIC